

(NASA-CR-195694) THE
TELECOMMUNICATIONS AND DATA
ACQUISITION REPORT Progress Report,
Jan. - Mar. 1991 (JPL) 160 p

N94-29613
--THRU--
N94-29626
Unclass

G3/32 0000683

The Telecommunications and Data Acquisition Progress Report 42-105

January-March 1991

E. C. Posner
Editor

May 15, 1991



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The Telecommunications and Data Acquisition Progress Report 42-105

January–March 1991

E. C. Posner
Editor

May 15, 1991



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

Preface

This quarterly publication provides archival reports on developments in programs managed by JPL's Office of Telecommunications and Data Acquisition (TDA). In space communications, radio navigation, radio science, and ground-based radio and radar astronomy, it reports on activities of the Deep Space Network (DSN) in planning, in supporting research and technology, in implementation, and in operations. Also included is standards activity at JPL for space data and information systems and reimbursable DSN work performed for other space agencies through NASA. The preceding work is all performed for NASA's Office of Space Operations (OSO). The TDA Office also performs work funded by two other NASA program offices through and with the cooperation of the Office of Space Operations. These are the Orbital Debris Radar Program (with the Office of Space Station) and 21st Century Communication Studies (with the Office of Aeronautics and Exploration Technology).

In the search for extraterrestrial intelligence (SETI), *The TDA Progress Report* reports on implementation and operations for searching the microwave spectrum. In solar system radar, it reports on the uses of the Goldstone Solar System Radar for scientific exploration of the planets, their rings and satellites, asteroids, and comets. In radio astronomy, the areas of support include spectroscopy, very long baseline interferometry, and astrometry. These three programs are performed for NASA's Office of Space Science and Applications (OSSA), with the Office of Space Operations funding DSN operational support.

Finally, tasks funded under the JPL Director's Discretionary Fund and the Caltech President's Fund that involve the TDA Office are included.

This and each succeeding issue of *The TDA Progress Report* will present material in some, but not necessarily all, of the following categories:

OSO Tasks:

DSN Advanced Systems

- Tracking and Ground-Based Navigation
- Communications, Spacecraft-Ground
- Station Control and System Technology
- Network Data Processing and Productivity

DSN Systems Implementation

- Capabilities for Existing Projects
- Capabilities for New Projects
- New Initiatives

- Network Upgrade and Sustaining

DSN Operations

- Network Operations and Operations Support
- Mission Interface and Support

- TDA Program Management and Analysis

- Ground Communications Implementation and Operations

- Data and Information Systems

- Flight-Ground Advanced Engineering

- Long-Range Program Planning

OSO Cooperative Tasks:

- Orbital Debris Radar Program

- 21st Century Communication Studies

OSSA Tasks:

Search for Extraterrestrial Intelligence

Goldstone Solar System Radar

Radio Astronomy

Discretionary Funded Tasks

Contents

OSO TASKS DSN Advanced Systems TRACKING AND GROUND-BASED NAVIGATION

Combining GPS and VLBI Earth-Rotation Data for Improved Universal Time.....	1
A. P. Freedman	
NASA Code 310-10-61-87-02	

COMMUNICATIONS, SPACECRAFT-GROUND

Computation of Vibration Mode Elastic-Rigid and Effective Weight Coefficients From Finite-Element Computer Program Output.....	13
R. Levy	
NASA Code BG310-20-65-86-04	
X-/Ka-Band Dichroic Plate Design and Grating Lobe Study.....	21
J. C. Chen	
NASA Code 310-20-64-89-01	
Site Comparison for Optical Visibility Statistics in Southern California.....	31
K. Cowles	
NASA Code 310-20-67-88-03	
Modifications of the Griesmer Bound.....	41
R. J. McEliece and G. Solomon	
NASA Code 310-30-71-83-02	
Integer Cosine Transform for Image Compression.....	45
K.-M. Cheung, F. Pollara, and M. Shahshahani	
NASA Code 310-30-71-83-02	

STATION CONTROL AND SYSTEM TECHNOLOGY

A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation, Part II: Tests on the Antenna.....	54
T. Y. Otoshi, S. R. Stewart, and M. M. Franco	
NASA Code 310-30-69-88-04	

DSN Systems Implementation CAPABILITIES FOR EXISTING PROJECTS

Phasing the Very Large Array on Galileo in the Presence of Jupiter's Strong Radio Emission.....	69
J. S. Ulvestad	
NASA Code 314-40-31-30-03	
Elevation Control System Model for the DSS 13 Antenna.....	83
W. Gawronski and J. A. Mellstrom	
NASA Code 314-40-42-10-40	
Radio Science Ground Data System for the Voyager-Neptune Encounter, Part I.....	109
E. R. Kursinski and S. Asmar	
NASA Code 314-40-41-10-13	

Capabilities for New Projects

Gaussian Beam and Physical Optics Iteration Technique for Wideband Beam Waveguide Feed Design	128
W. Veruttipong, J. C. Chen, and D. A. Bathker	
NASA Code 314-30-50-12-78	

Flight-Ground Advanced Engineering

Design and Analysis of a Low-Loss Linear Analog Phase Modulator for Deep Space Spacecraft X-Band Transponder Applications	136
N. R. Mysoor and R. O. Mueller	
NASA Code 314-30-55-30-01	

Orbital Debris Radar

Orbital Debris Radar Instrumentation	146
R. M. Goldstein and L. W. Randolph	
NASA Code 906-76-40-43-01	
Errata	149

1994025111

N94-29614

TDA Progress Report 42-105

442582

May 15, 1991

Combining GPS and VLBI Earth-Rotation Data for Improved Universal Time

A. P. Freedman

Tracking Systems and Applications Section

The Deep Space Network (DSN) routinely measures Earth orientation in support of spacecraft tracking and navigation using very long-baseline interferometry (VLBI) with the deep-space tracking antennas. The variability of the most unpredictable Earth-orientation component, Universal Time 1 (UT1), is a major factor in determining the frequency with which the DSN measurements must be made. The installation of advanced Global Positioning System (GPS) receivers at the DSN sites and elsewhere may soon permit routine measurements of UT1 variation with significantly less dependence on the deep-space tracking antennas than is currently required. GPS and VLBI data from the DSN may be combined to generate a precise UT1 series, while simultaneously reducing the time and effort the DSN must spend on platform-parameter calibrations. This combination is not straightforward, however, and a strategy for the optimal combination of these data is presented and evaluated. It appears that, with the aid of GPS, the frequency of required VLBI measurements of Earth orientation could drop from twice weekly to once per month. More stringent real-time Earth-orientation requirements possible in the future would demand significant improvements in both VLBI and GPS capabilities, however.

I. Introduction

One of the largest error sources for high-precision spacecraft tracking and navigation is Earth-orientation variability.^{1,2} The Deep Space Network (DSN) is currently engaged in monitoring platform-parameter variabil-

ity through the Time and Earth Motion Precision Observations (TEMPO) program of twice-weekly very long-baseline-interferometry (VLBI) measurements. This burden on the DSN radio telescopes is expected to increase as ever higher Earth-orientation accuracy is needed for space missions in the late 1990s and beyond.

Earth-orientation measurements can be made using a Global Positioning System (GPS) receiver network currently being installed at the three DSN sites and at a number of additional sites around the globe [1]. If an automated, rapid-turnaround data collection and reduction system is in place, these GPS data can be used in conjunc-

¹ R. Treuhaft and L. Wood, "Revisions in the Differential VLBI Error Budget and Applications for Navigation in Future Missions," JPL Interoffice Memorandum 335.4-601 (internal document), Jet Propulsion Laboratory, Pasadena, California, December 31, 1986.

² S. W. Thurman, "DSN Baseline Coordinate and Station Location Errors Induced by Earth Orientation Errors," JPL Engineering Memorandum 314-488 (internal document), Jet Propulsion Laboratory, Pasadena, California, July 25, 1990.

tion with DSN VLBI data to produce a highly accurate Earth-orientation series in near-real time [2]. Such a system would help to reduce the frequency with which VLBI Earth-orientation measurements must currently be made, thereby allowing more time for direct spacecraft tracking and communications.

The component of Earth orientation that is the most variable and unpredictable is the Earth's rotation rate, as measured by Universal Time 1 (UT1). Its unpredictability is a major motivation for the twice-weekly TEMPO measurements currently being performed. GPS techniques promise to deliver daily estimates of the UT1 rate of change (also known as length of day or LOD) with a precision of better than 0.1 msec/day [2]. Unfortunately, random error in LOD measurements can accumulate to produce estimates of UT1 (i.e., integrated LOD) that are in error by an amount that could exceed mission requirements for UT1 accuracy. (Currently, a 50-nrad tracking accuracy requirement necessitates knowledge of Earth rotation at the ~ 0.6 -msec level; future missions may require 0.1-msec-level real-time UT1 knowledge.³) Hence, some sort of synergism is required between the GPS and VLBI techniques such that VLBI data would constrain long-term excursions in the GPS-derived UT1 time series while the GPS data would guarantee accurate monitoring of high-frequency UT1 variability.

This article presents a method for combining DSN VLBI and GPS data. Section II discusses the nature of UT1 and LOD and highlights the need for frequent, accurate, LOD measurements. Section III reviews the capabilities of both VLBI and GPS for measuring UT1 and LOD. Section IV discusses the combination of these two data types and various issues that need to be considered in order to generate the most effective combination.

II. The Time Variability of UT1 and LOD

The angle of rotation of the Earth is represented by a time, UT1 (equal to the rotation angle times a constant), and is usually discussed with respect to a reference rotation angle that is growing at a constant, predefined rate. Two commonly used reference angles, both based on atomic time standards, are Universal Coordinated Time (UTC), which is always maintained within 1 second of UT1, and International Atomic Time (TAI). The LOD is a measure of the rate of change of UT1,⁴ and is defined by

$$L \approx -L_0 \frac{\partial U}{\partial t} \quad (1)$$

where L is the excess length of day, U represents UT1-UTC, and L_0 is the nominal length of the day (86,400 sec) [3]. LOD currently has a magnitude of about 2 msec, indicating that the value of UT1-UTC decreases by about 2 msec every 24 hours.

LOD can vary in an unpredictable manner, typically changing by a few hundredths of a millisecond over one day. Figure 1 shows a sample of smoothed LOD over a recent 1-year period; note both the stochastic, high-frequency behavior and the more pronounced regular variations at semi-annual and ~ 40 -day periods. In this and subsequent figures illustrating LOD, tidal effects have been removed according to the model of Yoder et al. [4].

If unmodeled, this variability in LOD can, within a few days, lead to an accumulated UT1 error that exceeds the current accuracy requirement of 0.6 msec. To illustrate this, Fig. 2 shows the growth in error in UT1-UTC as a function of time for three simple cases for the behavior of LOD, according to

$$U_{\text{error}} = U_{\text{true}} - U_{\text{predicted}} = \underbrace{\int_{t_0}^t [-L(\tau)] d\tau}_{U_{\text{true}}} - \underbrace{\int_{t_0}^t [-L_r] d\tau}_{U_{\text{predicted}}} \quad (2)$$

where $\tau \equiv \text{time}/L_0$ is dimensionless. In all three cases, the assumed reference value of UT1, $U_{\text{predicted}}$, is based on a value of LOD, L_r , that remains constant with time; thus, $U_{\text{predicted}}$ varies linearly with time. In case one, LOD itself grows linearly with time, i.e., $L(t) = L_r(t_0) + \alpha(t - t_0)$. In case two, LOD grows linearly with time for 2 days (as in case one), and then remains constant at its new value. In the third case, LOD is assumed to grow in proportion to the square root of time, i.e., $L(t) = L_r(t_0) + \alpha(t - t_0)^{1/2}$. The constant α is assumed to be 0.07 msec/day in the first two cases and 0.07 msec/ $\sqrt{\text{day}}$ in the third case. (Note that in these examples, the units of time for the integration are given in days, while LOD and UT1 are measured in milliseconds.) From Fig. 2, it is apparent that the maximum permissible UT1 error, 0.6 msec, is exceeded after 4 to 6 days. Although LOD behavior tends to be irregular, occasional secular variations in LOD (see Fig. 1) can result in unpredictable rapid growth of UT1 error akin to that illustrated by the three cases shown in Fig. 2; at times such as these, frequent measurements of LOD or UT1 are critical.

³ Treuhaft and Wood, op cit.

⁴ LOD as defined in this article is more accurately termed ΔLOD , the excess length of day (the difference between the length of the day and L_0), but the term LOD is retained based on common usage.

A more realistic model for the day-to-day variability of LOD and UT1 may be found from power spectral density (PSD) plots of LOD, such as that shown in Fig. 3. For periods longer than about 10 days, the spectrum is characterized by a $(\text{frequency})^{-2}$ dependence, corresponding to a random-walk stochastic process [5]. LOD can thus be characterized, over subseasonal time scales, as an integrated white-noise (i.e., a random-walk) process, while UT1 can be viewed as an integrated random walk. The white-noise PSD, Q , obtained from Fig. 3 is $Q \approx 0.005 \text{ msec}^2/\text{day}$. After a time $\Delta t = t - t_0$, the expected variance of LOD is $Q\Delta t$, while the corresponding expected error is $\sim 0.07\sqrt{\Delta t} \text{ msec}$ (when Δt is measured in days). This stochastic model for LOD still produces an unacceptably large UT1 error after 6 days.

Given the typical variability of LOD, accurate UT1 or LOD values must be available at least every 5 days in order to meet the current DSN requirements. In fact, the larger but less frequent LOD variations argue for measurements at least every 4 days. Furthermore, these time intervals do not take into account the required data processing time. The turnaround time for processing TEMPO data is typically two to three days, at which time the UT1 uncertainty has already grown to at least 0.2 msec. Since a new measurement must be available within 3 more days and it takes 2 to 3 days to process the next data point, a new TEMPO measurement must be made within one more day. The net effect, assuming a 2- or 3-day turnaround time, is to require UT1 measurements at least every 3 days.

III. Measuring UT1 and LOD With VLBI and GPS

Very long-baseline interferometry is sensitive to the orientation of the VLBI baseline in a celestial reference frame defined by quasars. Since the angle corresponding to UT1 is one component of the Earth's orientation in inertial space, UT1 is a directly measurable quantity, although multiple baselines are required to unambiguously remove the effects of polar motion. Twice-weekly TEMPO measurements, one per week on each of two different intercontinental baselines, provide regular UT1 and polar motion data to the DSN with typically a 2- or 3-day turnaround time (although 1-day turnaround has been demonstrated). Other VLBI networks, such as the National Geodetic Survey's International Radio Interferometric Surveying (IRIS) network and the U.S. Naval Observatory's VLBI network (NAVNET), also provide regular UT1 measurements, albeit with slower turnaround times that are unsuitable for the needs of the DSN. The formal errors listed for the TEMPO data are currently at the 0.1- to 0.2-msec level

[6], although combining and smoothing data from the two baselines can generate improved UT1 estimates with formal errors below 0.1 msec.

Each VLBI network relies on its own reference frames, both celestial and terrestrial, to produce Earth-rotation measurements. These reference frames may have small offsets with respect to one another, resulting in systematic differences between the Earth-rotation time series generated by each network [7]. These UT1 offsets are generally small and slowly varying, but they need to be considered in any precise estimate of UT1, especially if data from multiple techniques are combined.

GPS methods differ fundamentally from VLBI techniques. To determine Earth orientation with GPS, one first needs to define a terrestrial reference frame. This is usually done by fixing the locations of a few select ground receivers. These sites are known as fiducial sites and are tied by local ground surveys to nearby, collocated VLBI antennas whose relative positions are known precisely through VLBI [8]. A precise set of initial values, derived from VLBI and from satellite laser ranging (SLR), is used to orient this Earth-fixed frame with respect to a celestial reference frame, which then permits the GPS satellite orbit epoch states to be constrained in inertial space. The behavior of the satellite network is governed by dynamical models, such that movements of the solid Earth within that framework, i.e., Earth orientation, can be observed. UT1-UTC, however, cannot be determined simply by monitoring the relative motions of the terrestrial network and the GPS satellite orbits, as it is defined with respect to an arbitrary location on the celestial sphere. The signature in the data residuals of an error in UT1 is identical to that due to an error in the satellites ascending node locations. In addition, since the GPS satellites have almost identical orbital radii, inclinations, and eccentricities, the oblateness of the Earth induces precession of the GPS orbits that is nearly indistinguishable from a slow UT1 variation.

Although GPS techniques cannot directly measure the absolute value of UT1 and have difficulty distinguishing long-term UT1 changes from precession of the GPS orbits, they are sensitive to short-term changes in UT1. The large number of GPS satellites and receivers provide robust and redundant networks for measuring the displacements of the terrestrial frame with respect to the GPS orbits that result from rapid UT1 variations. Regular GPS measurements of LOD are, in fact, already being made, although with a less than ideal satellite configuration and station network [9]. These values are weekly averages and have typical errors of about 0.15 msec. The capability of GPS to measure LOD

is expected to improve dramatically by the mid-1990s. The full constellation of 24 satellites will be available, along with a worldwide, high-precision GPS tracking network with the potential for near-real-time data processing. LOD measurements are thus expected at the 0.05-msec level or better in the next few years [2].

The terrestrial reference frame used in GPS processing, although tied to a VLBI terrestrial frame through the fiducial technique, is not necessarily identical to that of VLBI, since local site ties between VLBI and GPS antenna phase centers may contain errors. These site ties can also vary over time, due either to changes in antenna positions or to improvements in the local ground surveys. In addition, as VLBI site location and site velocity estimates improve, both VLBI and GPS reference frames and the Earth orientation time series may be affected.

VLBI is a proven, currently operating technique to measure UT1, whereas precise, routine, operational determination of UT1 using GPS techniques lies at least a year or two in the future. Why, then, is there interest in supplementing VLBI with GPS?

As the tracking and navigation duties of the DSN expand and higher precision is needed, more frequent UT1 data will be required. As shown in Section II, the current frequency of TEMPO measurements is barely adequate in providing near-real-time Earth-rotation information; in fact, the TEMPO data must be supplemented with meteorologically derived atmospheric angular-momentum information (which has been shown empirically to be highly correlated with LOD) to ensure reasonable accuracy of current UT1 estimates [10]. The higher-precision future Earth-rotation requirements of the DSN will require more frequent VLBI data if alternative techniques are not available. Dedicating DSN radio telescopes to obtaining Earth rotation at daily or 2-day intervals would heavily tax the already-overburdened DSN tracking schedule. Hence, future high-precision Earth-platform determination is critically dependent on new high-accuracy technologies with the potential for rapid data turnaround, such as GPS.

IV. Combining VLBI and GPS Data

Three main issues need to be addressed in order to accurately and reliably combine GPS and VLBI Earth-rotation data. They are (1) growth of UT1 error with time, (2) reference-frame definition and compatibility, and (3) systematic technique-dependent errors. They will be discussed, in turn, below.

A. Growth of Error in UT1

Any strategy for the synergistic use of GPS and VLBI Earth-rotation data to monitor UT1 must take into account error characteristics of both GPS and VLBI, along with the observed variability of Earth rotation. Since VLBI is able to measure UT1 directly whereas the strength of GPS lies in its ability to monitor LOD (the change from day to day in UT1), VLBI is needed to provide an initial estimate of UT1. Subsequent measurements of LOD by GPS can then permit the calculation of UT1 by daily integration. Since errors in UT1 will accumulate over time, VLBI data are also needed to calibrate the GPS-determined time series. This periodic VLBI calibration should overcome various problems related to long-term monitoring of Earth rotation with GPS, including the effects of orbit precession, data noise, satellite force model errors, etc.

Figure 4 schematically illustrates this plan for GPS and VLBI synergism. At some epoch, a VLBI observation is made, yielding a precise estimate (and formal error) of UT1. Daily measurements of LOD are then obtained by a rapid-turnaround global GPS network, and UT1 and its uncertainty determined by integrating these LOD values. At some point, the accumulating uncertainty in UT1 necessitates a new VLBI observation, which fixes the value of UT1 at the new epoch and constrains its uncertainty. Note that Fig. 4 illustrates only a near-real-time method for estimating UT1; a UT1 time series of uniformly high precision will still be available after a delay of 2 weeks or more.

To determine the frequency with which periodic VLBI measurements should be made, the sources of UT1 error must be evaluated. Formal errors for UT1 derived from TEMPO measurements are currently at the 0.1-msec level, although VLBI system improvements should reduce this uncertainty in the next few years. In this study, a formal VLBI-derived UT1 error of either 0.1 msec (conservative) or 0.05 msec (optimistic) and a 2-day turnaround time for processing these data are assumed. A UT1 measurement is assumed to represent the average value of UT1 over a few hours, obtained from back-to-back, 3-hour TEMPO VLBI sessions on each of two baselines. (Note that this differs from the current observing strategy of one 3-hour session on alternate baselines every few days.)

Based on extensive covariance analyses [2], the LOD error anticipated from GPS when an optimal satellite and station configuration is established is less than 0.05 msec. To be conservative in this analysis, and for ease of presentation, we assume GPS-derived LOD also to be accurate to either 0.1 msec or 0.05 msec. Note that the GPS observable is, in actuality, an estimate of the UT1 rate-of-change

over a 12- to 24-hour period, but this is essentially equivalent to the average LOD. GPS data are assumed to be generated either every 24 or 48 hours, and to be available after a 24-hour processing delay.

All error estimates mentioned are the one-sigma formal errors and are assumed to be Gaussian. In addition, no correlation is assumed between VLBI and GPS errors, even though technically there may be some correlation if any VLBI-derived parameters are used to initialize the GPS data processing scheme.

Figure 5 illustrates a few examples of UT1 error behavior. (The Appendix describes the mathematical model used to generate this and the subsequent figure.) Both the VLBI UT1 and daily GPS LOD measurements are assumed to have identical formal errors (either 0.10 or 0.05 msec). At $t = 0$ days, a VLBI measurement of UT1 is made. The reduced data point is available at $t = 2$ days; by this time, however, the real-time knowledge of UT1 based solely on VLBI will have degraded due to the stochastic nature of LOD. Fortunately, GPS data are being obtained each day, providing a realistic assessment on day 2 of the LOD for day 0 and day 1 (recall that there is a 24-hour delay for GPS data). Subsequently, GPS-derived LOD data with their formal errors control the growth of uncertainty in UT1. At some point (in this case, $t = 40$ days), a new VLBI measurement is made in response to the accumulating uncertainty in UT1. Again, it takes 2 days for the measurement to become available, and the cycle continues.

In the case where formal errors for VLBI UT1 and GPS LOD are both 0.1 msec, the 0.6-msec DSN threshold for UT1 is reached in 35 days. Thus, DSN VLBI platform-parameter measurements would be needed approximately monthly. If the measurement errors for the two techniques drop to the 0.05-msec level, VLBI measurements of UT1 might be needed only once every 3 or 4 months. In this latter case, the DSN could even collect GPS LOD data every other day and still free up the VLBI system for almost 2 months.

If, however, DSN requirements for near-real-time Earth rotation drop to the 0.2-msec level, the picture is not so pleasant. This requirement could not be met with 0.1-msec quality data. Only with high-quality (~ 0.05 -msec) data from both VLBI and GPS, and GPS LOD measurements obtained daily, could a 0.2-msec requirement be met without VLBI calibrations for any appreciable length of time. In this case, DSN VLBI measurements would be needed about every 2 weeks. This is still less than half as often as TEMPO currently oper-

ates, and the alternative, TEMPO-only option would require TEMPO measurements to be taken daily. To meet a 0.1-msec real-time UT1 requirement, even VLBI and GPS data with errors at the 0.05-msec level are inadequate. Measurement errors must be reduced to the 0.02-msec level or better, and DSN VLBI calibration measurements must be made at least weekly to meet this more stringent Earth-rotation requirement.

Figure 6 addresses a subtle issue concerning the near-real-time need for UT1. On any given day, the estimate of UT1 for the preceding day is an empirical value based on combined VLBI and GPS data. To generate an actual real-time estimate of UT1, this value for the preceding day must be adjusted for the behavior expected of LOD over the additional day, and the error estimate corrected accordingly. Two cases are illustrated for this theoretical behavior: a random walk and worst case linear growth. The random-walk model is simply the stochastic behavior of LOD described in Section II, while the worst case scenario corresponds to an unusually large daily change in LOD. These two cases correspond to growth in the uncertainty of LOD over one day of 0.07 msec and 0.10 msec, respectively.

Figure 6 shows the expected error for these two models if GPS data are taken either daily or, for purposes of comparison, every other day. A VLBI UT1 measurement with a 0.05-msec formal error is made on day 0 and becomes available on day 2. GPS LOD measurements with a 0.05-msec formal error are made either daily or on alternate, even-numbered days. After every LOD measurement, the uncertainty in LOD is assumed to grow until the next LOD measurement, according to one of three models: no growth, random-walk (stochastic) growth, or rapid linear ("maximum") growth. The alternate-day LOD data curves oscillate because the uncertainty in UT1 24 hours after a GPS measurement is made is smaller than the UT1 uncertainty 48 hours after the preceding measurement is made. Shown for comparison are the estimated uncertainties in UT1 if no GPS data were available to constrain the variability of LOD (the "VLBI only" curves).

When daily GPS LOD data are used to estimate the current value of UT1, the modeled behavior of the LOD does not significantly affect real-time UT1 estimates on the day after the last data are taken. If GPS data are obtained on alternate days, however, the estimated UT1 uncertainties are significantly affected by the assumed LOD behavior. In this case, rapid LOD variations can produce UT1 errors exceeding 0.2 msec very quickly. Hence, once-a-day estimates of LOD by GPS are still the preferred measurement frequency.

B. Reference-Frame Compatibility

As discussed in Section III, GPS and VLBI may generate estimates of Earth-rotation parameters using slightly different terrestrial reference frames. Each frame should be internally stable over time, because a principal source of reference-frame instability, global tectonic motions, is a slowly varying process. Hence, differences between the two reference frames should manifest themselves in the Earth-rotation data as a more-or-less constant bias and, perhaps, a rate (i.e., a secularly growing offset) between the two time series. These bias and rate terms may be estimated empirically, based on previously measured time series of UT1 and LOD from the two techniques. The terrestrial frames may also shift relative to each other in an unpredictable manner when fiducial sites are changed, antennas are moved, or improved local site ties between the GPS and VLBI antenna phase centers are determined.

As long as the site network and data reduction method for each technique remain the same, the biases and rates should not significantly vary. However, changes in the reference site network, collocated site ties, VLBI radio source catalog, or data processing software may generate UT1 time-series glitches that need to be closely monitored. These time-series offsets or slope changes might be evident only weeks or months after they commence, when a sufficient quantity of data of diverse types has accumulated to distinguish the time-series characteristics of individual techniques. Fortunately, reference-frame adjustments are not expected to occur often.

C. Systematic Errors

Related to reference-frame discrepancies are the effects of systematic errors in each technique, either in data collection or in data processing, that contaminate the resulting UT1 and LOD time series. Unlike reference-frame errors, these need not be of a long period. Systematic GPS-derived LOD errors have yet to be investigated, since the data themselves are not yet available at nearly the required frequency or precision. However, numerous orbit and baseline studies have isolated a number of possible systematic error sources for GPS techniques in general. These include orbit mismodeling, incorrect fiducial site ground ties, and mismodeled atmospheric propagation effects [11].

Isolating these systematic errors will require periodic campaigns of simultaneous GPS and VLBI Earth-rotation measurements.⁵ Intercomparing these time series and in-

corporating the results of other Earth orientation measurement techniques, such as IRIS and NAVNET VLBI, SLR, and lunar laser ranging (LLR), should help to pinpoint systematic errors. This type of analysis is ongoing for currently existing techniques [12].

V. Conclusions

Within the next few years, high-precision GPS techniques will be able to generate Earth-rotation values of great utility to the DSN. By combining these LOD data with VLBI-derived estimates of UT1, a high-quality, real-time UT1 time series could become available for critical DSN spacecraft navigation and tracking applications. By reducing the amount of antenna time the DSN must commit to Earth-platform-parameter estimation, use of GPS data would permit a greater percentage of radio telescope time to be available for spacecraft telemetry and tracking. This would be all the more valuable given the expected increase in demand on the DSN radio antennas from the large number of planned and ongoing interplanetary missions over the next decade.

Although a high-quality GPS receiver network will already be in place at the DSN sites and at numerous other locations around the globe in support of TOPEX/POSEIDON tracking, the DSN must also establish a mechanism for rapid GPS data collection and processing. A prototype system already exists for sub-daily orbit determination,⁶ but a fully operational system with 12-hour or less turnaround time and an explicit goal of measuring Earth orientation has yet to be implemented.

Estimates of predicted UT1 uncertainty suggest that, given a small but reasonable GPS measurement error, the frequency of TEMPO VLBI measurements can be reduced from twice weekly to monthly or less (although this once-a-month TEMPO session must include measurements on both baselines). If a future real-time Earth-rotation requirement drops to a third of its present value (i.e., to ≤ 0.2 msec), the only feasible way to fulfill this requirement, short of daily TEMPO measurements, is the synergistic combination of daily GPS LOD measurements and weekly or semimonthly VLBI UT1 measurements. Furthermore, to achieve a 0.1-msec real-time Earth-rotation capability, significant improvements in both VLBI and GPS techniques are also needed. Thus, both VLBI and GPS appear to be techniques essential for meeting DSN platform-parameter requirements in the coming years.

⁵ The first of these multiple-technique simultaneous-observing sessions will occur in 1991.

⁶ G. Blewitt, personal communication, Jet Propulsion Laboratory, Pasadena, California, 1990.

References

- [1] R. E. Neilan, W. G. Melbourne, and G. L. Mader, "The Development of a Global GPS Tracking System in Support of Space and Ground-Based GPS Programs," to be published in *GPS and Other Radio Tracking Systems* (proceedings of IAG Symposium 102, August 3-12, 1989, Edinburgh, Scotland), ed. Y. Bock and N. Leppard, New York: Springer-Verlag, 1989.
- [2] A. P. Freedman, "Determination of Earth Orientation Using the Global Positioning System," *TDA Progress Report 42-99*, vol. July-September 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 1-11, November 15, 1989.
- [3] K. Lambeck, *The Earth's Variable Rotation: Geophysical Causes and Consequences*, Cambridge, England: Cambridge University Press, 1980.
- [4] C. F. Yoder, J. G. Williams, and M. E. Parke, "Tidal Variations of Earth Rotation," *J. Geophys. Res.*, vol. 86, no. B2, pp. 881-891, February 10, 1981.
- [5] J. O. Dickey, S. L. Marcus, and J. A. Steppe, "An Investigation of the Earth's Angular Momentum Budget at High Frequencies" (abstract), *EOS Trans. Am. Geophys. Un.*, vol. 70, no. 43, p. 1055, October 24, 1989.
- [6] J. A. Steppe, S. H. Oliveau, and O. J. Sovers, "Earth Rotation Parameters From DSN VLBI: 1990," *Earth Orientation and Reference Frame Determinations, Atmospheric Excitation Functions, Up to 1989 (Annex to the International Earth Rotation Service (IERS) Annual Report for 1989)*, Paris, France: Central Bureau of IERS-Observatoire de Paris, June 1990.
- [7] Z. Altamimi, E. F. Arias, C. Boucher, and M. Feissel, "Earth Orientation Determinations: Some Tests of Consistency," in *Earth Rotation and Coordinate Reference Frames, IAG Symposium 105*, Edinburgh, Scotland, August 11, 1989.
- [8] J. M. Davidson, C. L. Thornton, S. A. Stephens, G. Blewitt, S. M. Lichten, O. J. Sovers, P. M. Kroger, L. L. Skrumeda, J. S. Border, R. E. Neilan, C. J. Vegos, B. G. Williams, J. T. Freymueller, T. H. Dixon, and W. G. Melbourne, *The Spring 1985 High Precision Baseline Test of the JPL GPS-Based Geodetic System: A Final Report*, JPL Publication 87-35, Jet Propulsion Laboratory, Pasadena, California, November 15, 1987.
- [9] E. R. Swift, "Earth Orientation From GPS," presented at the Fifth Annual GPS Workshop, Jet Propulsion Laboratory, Pasadena, California, April 2, 1990.
- [10] A. P. Freedman, J. O. Dickey, J. A. Steppe, L.-Y. Sung, and T. M. Eubanks, "The Short-Term Prediction of Universal Time and Length-of-Day Using Atmospheric Angular Momentum," in preparation for *J. Geophys. Res.*, 1991.
- [11] S. M. Lichten and W. I. Bertiger, "Demonstration of Sub-Meter GPS Orbit Determination and 1.5 Parts in 10^8 Three-Dimensional Baseline Accuracy," *Bulletin Geodesique*, vol. 63, pp. 167-189, 1989.
- [12] IERS, *International Earth Rotation Service (IERS) Annual Report for 1989*, Paris, France: Central Bureau of IERS-Observatoire de Paris, 1990.

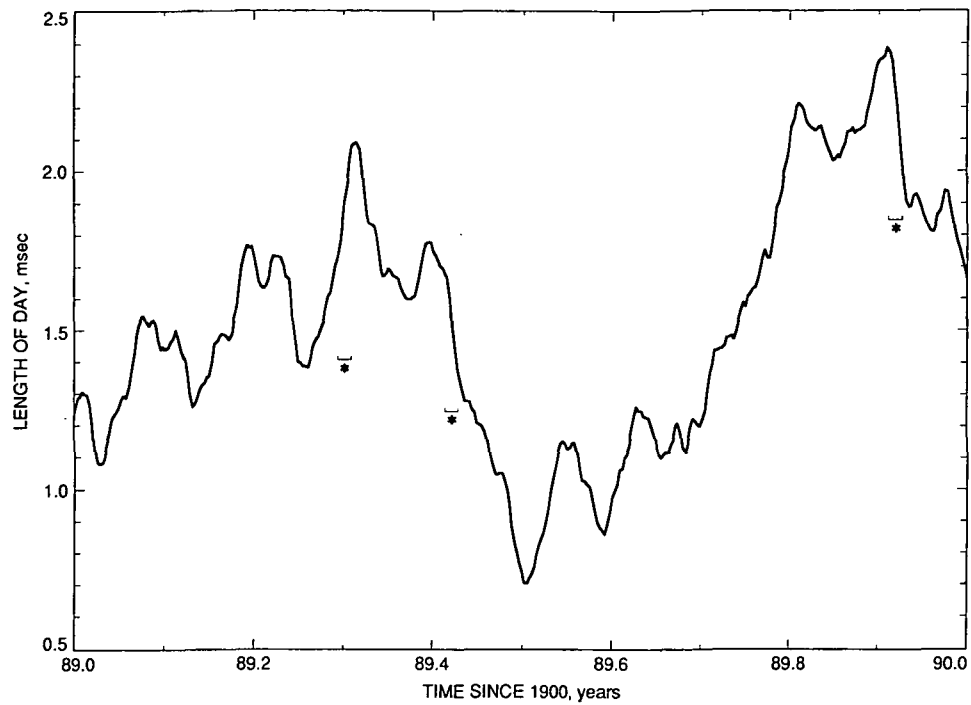


Fig. 1. Sample LOD time series. The asterisks mark selected time periods when LOD changed unusually rapidly.

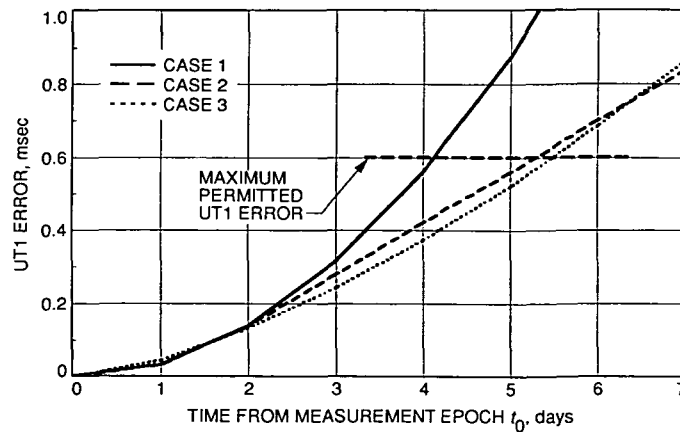


Fig. 2. Growth in UT1-UTC error as a function of time for three different LOD behaviors.

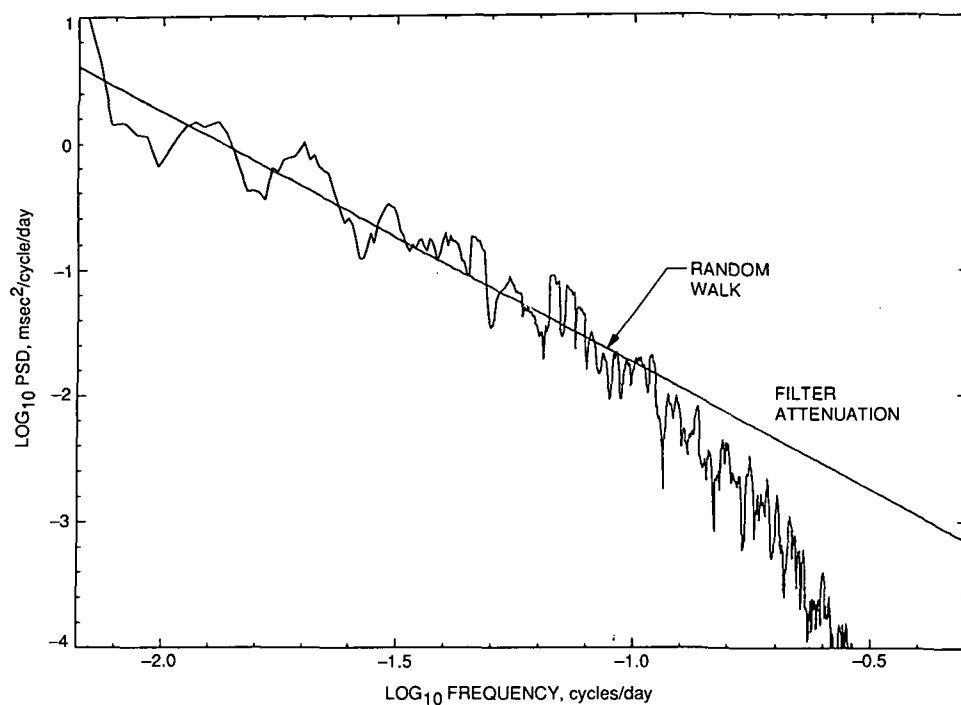


Fig. 3. PSD of smoothed multiyear LOD time series (after [5]).

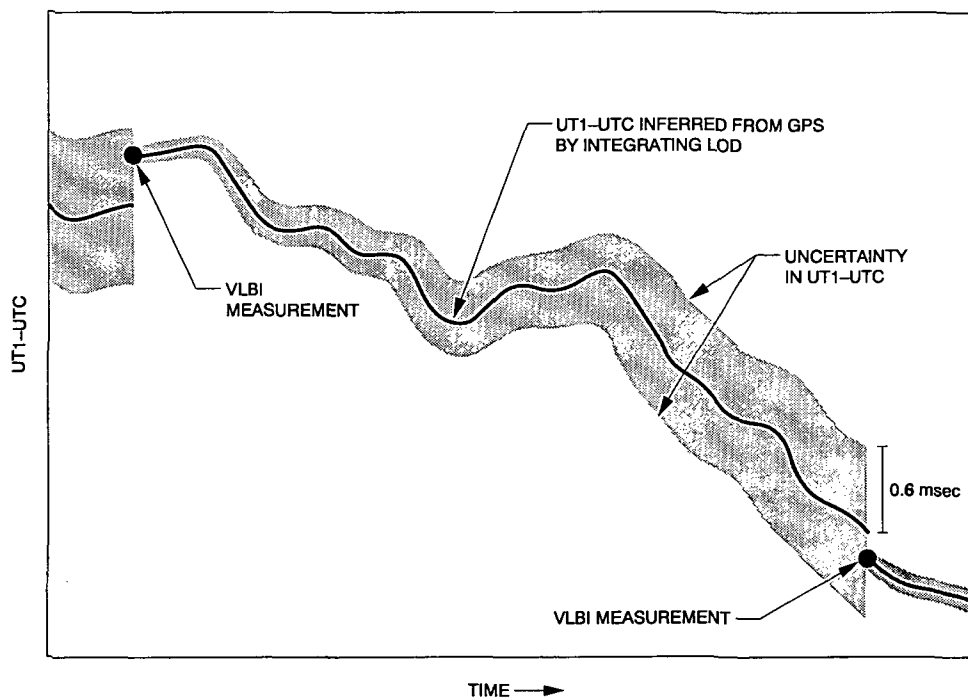


Fig. 4. The strategy for combining GPS and VLBI observations to achieve a high-precision, near-real-time UT1 time series (after [2]).

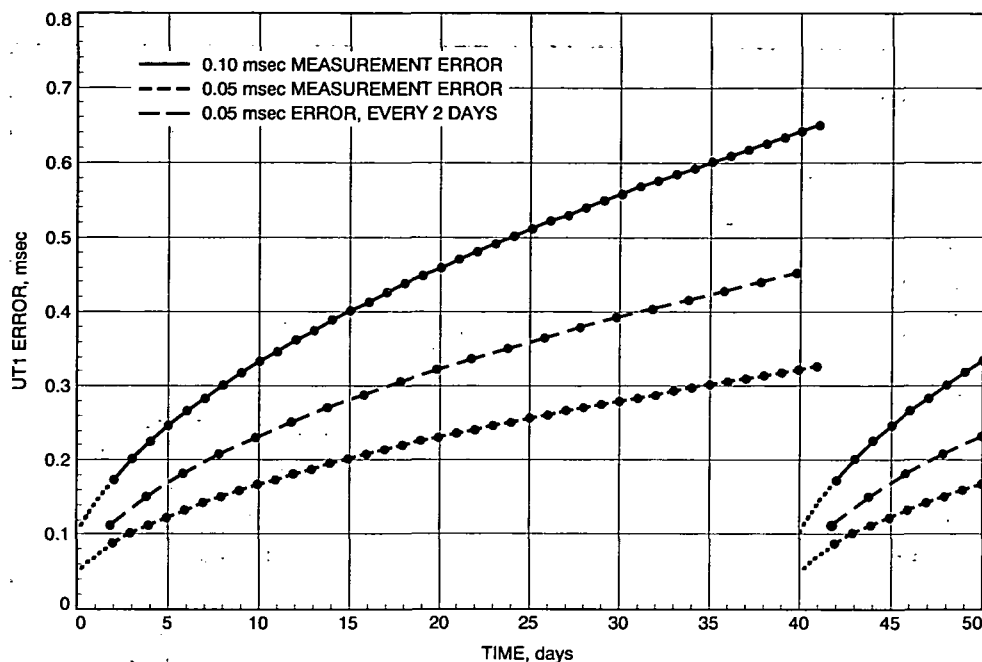


Fig. 5. Estimated one-sigma error in real-time UT1-UTC as a function of time for three different measurement scenarios. Lines are used to connect the points only; they do not represent the estimated errors at intermediate times. Dotted lines indicate retroactive uncertainty in UT1 following the VLBI measurement.

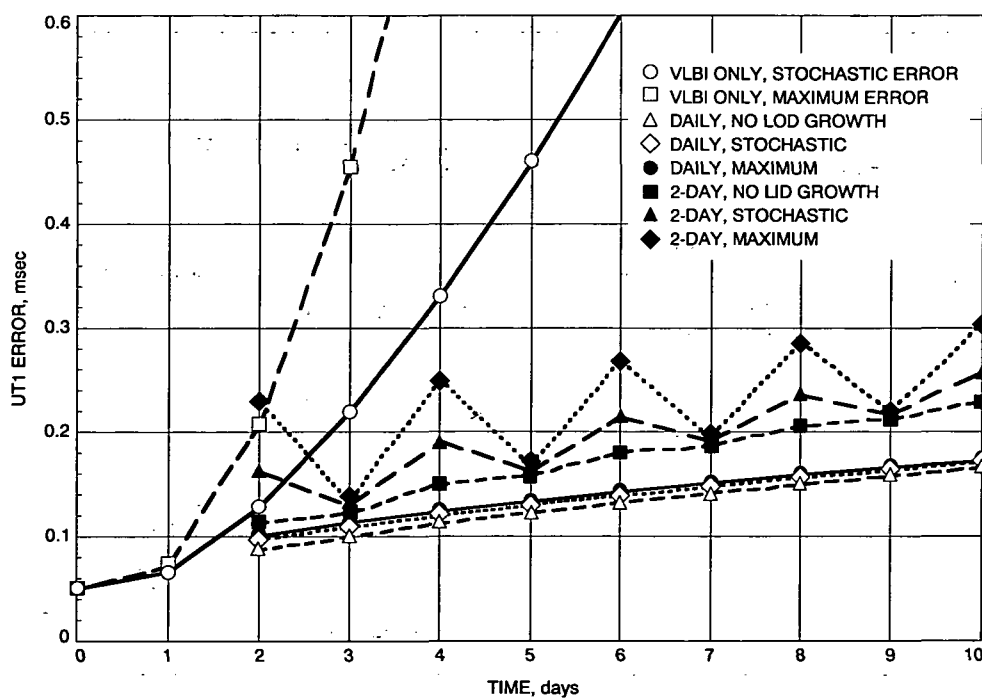


Fig. 6. Expected real-time one-sigma error in UT1-UTC when an LOD behavior model is included. Note that lines are used to connect the points only; they do not represent the estimated errors at intermediate times.

Appendix

Modeling the UT1 Error Structure

The mathematical model for the evolution of UT1 error estimates shown in Figs. 5 and 6 is presented below. Measured data are assumed to consist of the following: $U_{m|0}$ (UT1-UTC at time t_0) with formal error $\sigma_{U|0}$, and LOD estimates ($\bar{L}_{m|i}$) at times $t_i, i = 0, \dots, l$, with formal errors $\sigma_{L|i}$. The true measurement errors in UT1 and LOD are denoted by $\epsilon_{U|0}$ and $\epsilon_{L|i}$, respectively. All times are normalized, i.e., $t_i = \text{time}(i)/L_0$. Recall that LOD estimates are actually estimates of $-\Delta UT1/\Delta t$ over a time interval Δt .

The predicted value of UT1 at t_n , where $t_n > t_l$, is

$$U_p(t_n) = U_{m|0} - \sum_{i=1}^l \bar{L}_{m|i-1} (t_i - t_{i-1}) - \int_{t_l}^{t_n} L'(\tau) d\tau \quad (\text{A-1})$$

where $L'(\tau)$ is the predicted behavior of LOD at normalized time τ . In the absence of data, LOD is usually assumed to remain constant, i.e., $L'(\tau) = \bar{L}_{m|l}$ for $\tau \geq t_l$. The true value of UT1 is represented by

$$U(t_n) = U_0 - \int_{t_0}^{t_n} L(\tau) d\tau = [U_{m|0} + \epsilon_{U|0}] - \left[\int_{t_0}^{t_l} L(\tau) d\tau + \int_{t_l}^{t_n} L(\tau) d\tau \right] \quad (\text{A-2})$$

where U_0 and $L(\tau)$ denote the true UT1 and LOD values, respectively.

The true value of LOD may be modeled within the above equation as

$$\int_{t_0}^{t_l} L(\tau) d\tau = \sum_{i=1}^l \int_{t_{i-1}}^{t_i} L(\tau) d\tau = \sum_{i=1}^l \bar{L}_{m|i-1} (t_i - t_{i-1}) \quad (\text{A-3})$$

where $\bar{L}_{m|i-1}$ is the mean value of LOD over the interval t_{i-1} to t_i ; it is assumed that $\bar{L}_{m|i-1} = \bar{L}_{m|i-1} + \epsilon_{L|i-1}$. The final term in Eq. (A-2), the effect of LOD variability following the last LOD measurement at time t_l , is modeled as

$$\int_{t_l}^{t_n} L(\tau) d\tau = \int_{t_l}^{t_n} [L'(\tau) + \eta(\tau)] d\tau = \int_{t_l}^{t_n} [\bar{L}_{m|l} + \eta(\tau)] d\tau \quad (\text{A-4})$$

where the difference between the true and assumed LOD is

$$\eta(\tau) = \epsilon_{L|l} + \int_{t_l}^{\tau} \omega(\mu) d\mu \quad (\text{A-5})$$

In Eq. (A-5), the first term on the right is the LOD measurement error at t_l , while the final term in this expression emerges from the random-walk nature of LOD. The variable $\omega(\mu)$ describes a white-noise stochastic process whose integral over time generates the random-walk variations of LOD. Alternatively, the true LOD may be modeled as an analytic function of time, i.e., $\eta(\tau) = \epsilon_{L|l} + f(\tau - t_l)$.

The error terms and white-noise behavior in the above equations may be described by the following set of expectation values:

$$\left. \begin{aligned} \langle \varepsilon_{U|0} \varepsilon_{U|0} \rangle &= (\sigma_{U|0})^2 \\ \langle \varepsilon_{L|i} \varepsilon_{L|j} \rangle &= (\sigma_{L|i})^2 \delta_{ij} \\ \langle \omega(\mu) \omega(\rho) \rangle &= Q \delta(\mu - \rho) \end{aligned} \right\} \quad (\text{A-6})$$

while all other expectation values of $\omega(\mu)$, $\varepsilon_{U|0}$, $\varepsilon_{L|i}$ and their products are assumed to be zero. In Eq. (A-6), δ_{ij} is the Kronecker delta, $\delta(\mu - \rho)$ is the Dirac delta function, and Q is the PSD of the white-noise process.

The error in UT1 as a function of time is given by

$$\varepsilon_U(t_n) \equiv U_p(t_n) - U(t_n) = -\varepsilon_{U|0} + \sum_{i=1}^l [\varepsilon_{L|i-1}(t_i - t_{i-1})] + \int_{t_l}^{t_n} \eta(\tau) d\tau \quad (\text{A-7})$$

The variance of this error is

$$\begin{aligned} \text{Var}(\varepsilon_U(t_n)) &= \langle \varepsilon_U(t_n) \varepsilon_U(t_n) \rangle \\ &= (\sigma_{U|0})^2 + \sum_{i=1}^l (\sigma_{L|i-1})^2 (t_i - t_{i-1})^2 + \int_{\tau=t_l}^{t_n} \int_{\rho=t_l}^{t_n} \langle \eta(\tau) \eta(\rho) \rangle d\rho d\tau \end{aligned} \quad (\text{A-8})$$

where most of the expectation values have been evaluated. The rightmost term in Eq. (A-8) consists of two parts, one corresponding to the error of the final LOD measurement, $\varepsilon_{L|l}$, and one corresponding to either the random-walk or analytic model for LOD. Thus

$$\int_{\tau=t_l}^{t_n} \int_{\rho=t_l}^{t_n} \langle \eta(\tau) \eta(\rho) \rangle d\rho d\tau = (\sigma_{L|l})^2 (t_n - t_l)^2 + \left\{ \begin{aligned} &\frac{Q(t_n - t_l)^3}{3} \\ &\left[\int_{t_l}^{t_n} f(\tau - t_l) d\tau \right]^2 \end{aligned} \right. \quad (\text{A-9})$$

The upper quantity on the right corresponds to the random-walk generated uncertainty, and the lower to the uncertainty due to the modeled analytic behavior of LOD. Figure 5 is generated using the first two terms on the right-hand side of Eq. (A-8), while Fig. 6 utilizes all of Eqs. (A-8) and (A-9).

1994025112

N94-29615

TDA Progress Report 42-105

May 15, 1991

442583

Computation of Vibration Mode Elastic-Rigid and Effective Weight Coefficients From Finite-Element Computer Program Output

R. Levy

Ground Antennas and Facilities Engineering Section

Post-processing algorithms are given to compute the vibratory elastic-rigid coupling matrices and the modal contributions to the rigid-body mass matrices and to the effective modal inertias and masses. Recomputation of the elastic-rigid coupling matrices for a change in origin is also described. A computational example is included. The algorithms can all be executed by using standard finite-element program eigenvalue analysis output with no changes to existing code or source programs.

I. Introduction

The concepts of effective vibratory modal mass [1] and the application to the combined structural-mechanical-control system model [2] are important tools in the analysis and simulation [3] of the dynamic response of antennas and other complex structures with stringent performance or safety [4] requirements. Effective modal masses or inertias and their components associated with particular coordinate axes also provide the analyst with insight into the characteristics of particular vibratory modes. Simplifications of the transient dynamic analysis procedures can be facilitated by the convenience of identifying modes with relatively insignificant masses or inertias associated with the coordinate axis of interest. Such modes are candidates for modal truncation and elimination from the analytical model and replacement by contributions to residual masses or inertias.

The following discussion will show how the key matrix to this development, the elastic-rigid coupling ma-

trix, can be extracted from finite-element program eigenvalue analysis output, such as from the JPL-IDEAS [5], NASTRAN [6], or other typical finite-element analysis programs. Generation of the modal effective inertia matrix from the elastic-rigid coupling matrix follows readily. Finally, it will be shown how the effective inertia matrix with respect to one reference origin can be modified to relate a different origin by after-the-fact computations. Explicit example calculations are included.

II. Rigid-Body Mass Matrix and Modal Contributions

A finite-element structure model with N nodes and three translational and three rotational degrees of freedom at each node has a rigid-body transformation matrix, ϕR , with $6N$ rows and 6 columns. The columns refer to the three rigid-body translational displacements in a Cartesian X-Y-Z coordinate system and the three corresponding rigid-body rotations about these axes with respect to

a selected reference point origin. That is, each set of six rows of ϕR has the form:

$$\text{typical nodal sextet of rows of } \phi R = \begin{bmatrix} 1 & 0 & 0 & 0 & dZ & -dY \\ 0 & 1 & 0 & -dZ & 0 & dX \\ 0 & 0 & 1 & dY & -dX & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

In the above, dX , dY , and dZ are the differences between the coordinates of the particular node and the reference point. Also, if, as in the JPL-IDEAS program, the finite-element model does not contain rotational degrees of freedom, only the first three rows are applicable and ϕR is a $3N$ -by-6 matrix.

The elastic-rigid coupling matrix computation involves the finite-element symmetrical physical mass matrix, M , and the eigenvector matrix, ϕN . With three degrees of freedom per node, the mass matrix size is $3N$ by $3N$, and in the special case when the complete eigenvector matrix is available, this matrix is also the same size as the mass matrix. The elastic-rigid coupling matrix, MER , is computed as

$$MER = \phi N^t M \phi R \quad (2)$$

Note that MER has either $3N$ or $6N$ rows and 6 columns.

A natural-vibration-mode generalized mass matrix, MNN , is diagonal (from orthogonality of the normal modes) and is computed from the eigenvectors and mass matrix as

$$MNN = \phi N^t M \phi N \quad (3)$$

A rigid-body mass matrix, MRR , is ordinarily computed from the coordinate geometry and the mass matrix. However, it can also be computed from the previously mentioned expressions as

$$MRR = MER^t MNN^{-1} MER \quad (4)$$

To show this, the eigenvector matrix must be complete. Therefore, the inverse exists and substituting Eqs. (2) and (3) in Eq. (4) leads to

$$MRR = \phi R^t M \phi R \quad (5)$$

which is the rigid-body mass matrix by definition.

The rigid-body mass matrix is square, with six rows and columns. It is informative to partition MRR into four 3-by-3 matrices. If this is done, the upper left partition is diagonal and contains the sums of the nodal masses (usually all identical) in the coordinate axes' directions. The upper right partition (and its transpose in the lower left) contains the static (first) moments of the mass. Finally, the lower right partition is the inertia tensor, which has the mass moments of inertia about the respective coordinate axes on the diagonal and the products of inertia as off-diagonals. All dimensions implied in the computed terms of this matrix are with respect to the reference point.

Usually not all columns of the eigenvector matrix are available, or else, although available, some (or many) columns are truncated to condense the solution and calculations. The previous procedures will be applied to a particular retained column of eigenvectors. In the following, the subscript " j " refers to terms associated with a particular retained j th natural vibration mode. Thus, the j th mode elastic-rigid coupling matrix is the 1-by-6 row matrix MER_j ,

$$MER_j = \phi N_j^t M \phi R \quad (6)$$

and the contribution to the rigid-body mass matrix is the 6-by-6 matrix MRR_j

$$MRR_j = \frac{MER_j^t MER_j}{MNN_j} \quad (7)$$

In MRR_j above, the diagonals are the effective modal masses (1 to 3) and mass moments of inertia (4 to 6), which, when multiplied by the acceleration of gravity, are equivalent to the "Reduction to the Diagonals of the Residual Weight Matrix" in the JPL-IDEAS output. As in the discussion of the rigid-body mass matrix, the lower right-hand 3-by-3 partition is the contribution of this natural mode to the inertia tensor. If all these MRR_j matrices were computed and summed, the sum would (to within computer accuracy) be equal to the rigid-body mass matrix of Eq. (5).

III. Recovery of Elastic-Rigid Modal Coupling Matrices

In the JPL-IDEAS program formulation, the term "weight" is used in place of "mass" and "inertia," and implies weight moment of inertia, rather than mass moment of inertia. The conversion to the conventional mass formulation is made by dividing the respective weight terms by the acceleration of gravity. JPL-IDEAS directly provides a table of the elastic-rigid coupling weight matrix for each natural frequency mode under the heading "Un-Normalized Rigid Elastic Coupling Matrix." The six columns of the row matrix associated with each mode are tabulated under the headings "SUMX SUMY SUMZ SUMTX SUMTY SUMTZ." The first three headings refer to translational terms in the Cartesian axes' directions, and the last three refer to rotational terms about these axes.

In general, the elastic-rigid modal coupling matrices are not part of the usual output of finite-element programs. However, they can be recovered by post-processing the modal reactions. The computations use the j th mode reaction vector, \mathbf{R}_j , and the j th mode circular frequency, ω_j . Also, it is necessary to construct a rigid-body transformation matrix $\phi\mathbf{R}\mathbf{R}$ for the reaction points. This matrix has the same format as $\phi\mathbf{R}$ in Eq. (1), except that the differences in coordinates are between the reaction points and the reference point. Consequently, from statics, the j th mode reaction forces produce the following set of forces and moments at the reference point:

$$\mathbf{FMR}_j = \phi\mathbf{R}\mathbf{R}^t \mathbf{R}_j \quad (8)$$

On the other hand, forces and moments at the reference point for a natural mode of vibration can also be computed directly from the masses at the nodal points. That is, if \mathbf{F}_j are the forces of acceleration in the j th vibration mode, then

$$\mathbf{F}_j = -\omega_j^2 \mathbf{M}\phi\mathbf{N}_j \quad (9)$$

At the reference point, these forces produce a force and moment vector \mathbf{FMN}_j , which from statics is:

$$\mathbf{FMN}_j = \phi\mathbf{R}^t \mathbf{F}_j \quad (10)$$

Then, setting Eqs. (8) and (10) equal, using Eq. (9) and transposing, then using Eq. (6) and solving for \mathbf{MER}_j shows

$$\mathbf{MER}_j = \frac{\mathbf{R}_j^t \phi\mathbf{R}\mathbf{R}}{\omega_j^2} \quad (11)$$

which is the algorithm that can be used to compute the elastic-rigid coupling matrix.

IV. Changing the Reference Origin

If, after completing the previously described computations, there is the need to use an alternative reference origin, it is not necessary to repeat all these computations. Naturally, it is possible to repeat the procedure to recover the \mathbf{MER}_j from the reactions by changing $\phi\mathbf{R}\mathbf{R}$ accordingly and repeating the computations of Eq. (11). However, if there are many components in the reaction matrix, the following alternative could prove to be simpler.

To illustrate, let dox , doy , and doz be the change in coordinates of the reference point, and $\mathbf{D}\phi\mathbf{R}\mathbf{R}$ be the change in $\phi\mathbf{R}\mathbf{R}$ corresponding to dox , doy , and doz . Then, from Eq. (11), the change in the elastic-rigid coupling matrix is

$$\mathbf{DMER}_j = \frac{-\mathbf{R}_j^t \mathbf{D}\phi\mathbf{R}\mathbf{R}}{\omega_j^2} \quad (12)$$

It is evident from Eq. (1) that $\mathbf{D}\phi\mathbf{R}\mathbf{R}$ will be null except for the last three columns of the first three rows, which will have the form:

$$\begin{array}{c} \text{typical nodal triad} \\ \text{for first three rows} \\ \text{and last three} \\ \text{columns of } \phi\mathbf{R} \end{array} = \begin{array}{ccc} \begin{matrix} (4) & (5) & (6) \end{matrix} \\ \left[\begin{array}{ccc} 0 & doz & -doy \\ doz & 0 & dox \\ doy & -dox & 0 \end{array} \right] \end{array} \quad (13)$$

Therefore, it can be seen that only the fourth, fifth, and sixth columns of \mathbf{MER}_j will be affected. Concentrating on the change in the fourth column of \mathbf{DMER}_j , Eq. (13) can be rewritten as

$$\begin{array}{c} \text{nodal change triad to} \\ \text{compute column 4} \\ \text{of } \mathbf{DMER} \end{array} = \left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array} \right] \left\{ \begin{array}{c} doz \\ doy \\ -doz \end{array} \right\} \quad (14)$$

From Eqs. (12), (14), and the form of Eq. (1), and by identifying the components of \mathbf{MER} as the row vector

$$\text{MER}_j = [C1 \ C2 \ C3 \ C4 \ C5 \ C6]_j \quad (15)$$

it can be seen that the fourth column of DMER_j can be represented as

$$\text{DMER}_j(\text{column 4}) = [0 \ -C3 \ C2] \begin{Bmatrix} \text{dox} \\ \text{doy} \\ \text{doz} \end{Bmatrix} \quad (16)$$

By repeating a similar procedure for the fifth and sixth columns and transposing, the result is

$$\text{DMER}_j(\text{columns 4, 5, 6}) = [\text{dox} \ \text{doy} \ \text{doz}] C_j \quad (17)$$

where

$$C_j = \begin{bmatrix} 0 & C3 & -C2 \\ -C3 & 0 & C1 \\ C2 & -C1 & 0 \end{bmatrix}_j \quad (18)$$

V. Example Problem Computations

This example will illustrate computation of the elastic-rigid coupling matrices, the modal contributions to the rigid-body mass matrix and effective masses and inertias, and the computation of the changes in the elastic-rigid coupling matrix for a change in the reference origin.

The example structure is completely restrained in the X , Y , and Z coordinate directions at three foundation nodes. The analytical model represents only translational degrees of freedom, so that $\phi \mathbf{R}$ and $\phi \mathbf{R} \mathbf{R}$ have three rows for each of their associated nodes. The computations will cover the first three natural modes.

Table 1 contains the needed information supplied by the finite-element program. This consists of the coordinates of the foundation nodes, the reference origin, and partial eigenvalue analysis output. The latter consists of the frequencies and generalized masses and the reactions for the first three natural modes. Note that for this problem, no information other than the eigenvalue analysis frequencies, generalized masses, and reactions is needed about the remainder of the finite-element model or its response.

Table 2 shows the solutions for the modal contributions to the rigid-body mass matrices. The diagonal elements of these 6-by-6 matrices are the effective masses and inertias. These provide a convenient, although partially qualitative, means to characterize the nature of the vibration mode. For example, in the first natural mode there will be almost no mass moving in the X and Z directions (masses = $4.5\text{e}-04$ and $1.49\text{e}-02$), nor moment of inertia with respect to rotations about the Y axis (inertia = $7.9\text{e}+01$). The strong motion will be in the Y direction (mass = $4.0\text{e}+0$) with rotations about the X and Z axes (inertias = $3.91\text{e}+04$ and $1.37\text{e}+04$).

Table 3 illustrates the computations for the changes in elastic-rigid coupling matrices when the reference origin is changed. Here it can be seen that a change in only the Z -coordinate of the origin produces changes only in the rotational terms associated with the X and Z axes.

VI. Summary

Recovery of the elastic-rigid coupling matrices and computations for the modal contributions to the rigid-body mass matrix and effective modal masses and inertias have been described. A method to recompute the elastic rigid coupling matrix for a change in origin is also given. A computational example is included to demonstrate the procedures. All the described computations are performed by post-processing of standard finite-element program eigenvalue analysis output and no modifications or code changes in existing programs are required.

References

- [1] B. K. Wada, R. Bamford, and J. A. Garba, "Equivalent Spring-Mass System: A Physical Interpretation," *The Shock and Vibration Bulletin*, No. 42, Part 5, pp. 215-225, January 1972.
- [2] R. Levy, "Condensed Antenna Structural Models for Dynamics Analysis," *TDA Progress Report 42-80*, vol. October-December 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 40-61, February 15, 1985.
- [3] R. E. Hill, "A New State Space Model for the NASA/JPL 70-Meter Antenna Servo Controls," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 247-264, November 15, 1987.
- [4] R. W. Clough, "Earthquake Response of Structures," in *Earthquake Engineering*, edited by R. L. Wiegel, New Jersey: Prentice-Hall, pp. 307-334, 1970.
- [5] R. Levy and D. Strain, *JPL-IDEAS, Iterative Design of Antenna Structures*, COSMIC, NPO 17783, University of Georgia, Athens, Georgia, October 1988.
- [6] *The NASTRAN Theoretical Manual*, COSMIC, NASA SP 221(06) University of Georgia, Athens, Georgia, January 1981.

Table 1. Example Input data

Coordinates of the restrained nodes				
	<i>X</i>	<i>Y</i>	<i>Z</i>	
	50	0	30	
	-50	0	0	
	0	100	-20	
Coordinates of the reference point				
	<i>X</i>	<i>Y</i>	<i>Z</i>	
	0	0	50	
Eigenvalue analysis results				
Mode		1	2	3
<i>w</i>		1.1920e+02	1.6000e+02	2.8570e+02
MNN		3.9327e+00	5.8179e+00	2.8151e+00
Reaction node 1	<i>X</i>	3.5228e+04	-8.8273e+04	-1.9411e+03
	<i>Y</i>	5.9308e+02	3.4942e+04	-8.4705e+04
	<i>Z</i>	-2.9574e+04	1.5995e+05	-1.9594e+05
Reaction node 2	<i>X</i>	-6.5662e+03	-7.4353e+04	-2.5214e+03
	<i>Y</i>	-6.9630e+03	-3.3808e+04	2.8558e+03
	<i>Z</i>	-6.8021e+04	-2.1592e+05	4.1937e+04
Reaction node 3	<i>X</i>	-2.9258e+04	2.5865e+04	7.9293e+04
	<i>Y</i>	-5.0125e+04	-1.8745e+04	-1.2541e+05
	<i>Z</i>	9.4154e+04	3.0557e+04	2.1629e+05

Table 2. Solutions for modal contributions

Relative coordinates of the restrained nodes								
		X	Y	Z				
		50	0	-20				
		-50	0	-50				
		0	100	-70				
$\phi_{RR} =$ [Eq. (1)]	Node 1	1	0	0	0	-20	0	
		0	1	0	20	0	50	
		0	0	1	0	-50	0	
	Node 2	1	0	0	0	-50	0	
		0	1	0	50	0	-50	
		0	0	1	0	50	0	
	Node 3	1	0	0	0	-70	-100	
		0	1	0	70	0	0	
		0	0	1	100	0	0	
	Columns of MER, Eq. (11)							
	Mode	C1	C2	C3	C4	C5	C6	
1	4.1960e-02	3.9761e+00	2.4218e-01	-3.9204e+02	1.7633e+01	-2.3251e+02		
2	5.3422e+00	6.8793e-01	9.9270e-01	-2.9375e+01	5.9066e+02	-3.3242e+01		
3	-9.1677e-01	2.5392e+00	-7.6309e-01	-1.3843e+02	-7.9734e+01	1.5078e+02		
Modal contributions to rigid body mass matrix, Eq. (7)								
$MRR_1 =$	4.4771e-04	4.2424e-02	2.5840e-03	-4.1830e+00	1.8814e-01	-2.4808e+00		
	4.2424e-02	4.0200e+00	2.4485e-01	-3.9637e+02	1.7828e+01	-2.3507e+02		
	2.5840e-03	2.4485e-01	1.4913e-02	-2.4142e+01	1.0859e+00	-1.4318e+01		
	-4.1830e+00	-3.9637e+02	-2.4142e+01	3.9082e+04	-1.7578e+03	2.3178e+04		
	1.8814e-01	1.7828e+01	1.0859e+00	-1.7578e+03	7.9061e+01	-1.0425e+03		
	-2.4808e+00	-2.3507e+02	-1.4318e+01	2.3178e+04	-1.0425e+03	1.3746e+04		
$MRR_2 =$	4.9055e+00	6.3169e-01	9.1154e-01	-2.6973e+01	5.4237e+02	-3.0524e+01		
	6.3169e-01	8.1344e-02	1.1738e-01	-3.4734e+00	6.9843e+01	-3.9307e+00		
	9.1154e-01	1.1738e-01	1.6938e-01	-5.0122e+00	1.0078e+02	-5.6721e+00		
	-2.6973e+01	-3.4734e+00	-5.0122e+00	1.4831e+02	-2.9823e+03	1.6784e+02		
	5.4237e+02	6.9843e+01	1.0078e+02	-2.9823e+03	5.9967e+04	-3.3749e+03		
	-3.0524e+01	-3.9307e+00	-5.6721e+00	1.6784e+02	-3.3749e+03	1.8994e+02		
$MRR_3 =$	2.9855e-01	-8.2690e-01	2.4851e-01	4.5079e+01	2.5966e+01	-4.9102e+01		
	-8.2690e-01	2.2903e+00	-6.8829e-01	-1.2486e+02	-7.1918e+01	1.3600e+02		
	2.4851e-01	-6.8829e-01	2.0685e-01	3.7523e+01	2.1613e+01	-4.0872e+01		
	4.5079e+01	-1.2486e+02	3.7523e+01	6.8067e+03	3.9207e+03	-7.4142e+03		
	2.5966e+01	-7.1918e+01	2.1613e+01	3.9207e+03	2.2583e+03	-4.2706e+03		
	-4.9102e+01	1.3600e+02	-4.0872e+01	-7.4142e+03	-4.2706e+03	8.0758e+03		

Table 3. Solutions for change of origin

Change in reference origin			
	X	Y	Z
Original	0	0	50
New	0	0	0
Change	0	0	-50

Mode 1			
Eq. (18) $C =$	$\begin{bmatrix} 0 & 2.4218e-01 & -3.9761e+00 \\ -2.4218e-01 & 0 & 4.1960e-02 \\ 3.9761e+00 & -4.1960e-02 & 0 \end{bmatrix}$		
Eq. (17) DMER =	[0 0 0]	-1.9880e+02	2.0980e+00 0]

Mode 2			
$C =$	$\begin{bmatrix} 0 & 9.9270e-01 & -6.8793e-01 \\ -9.9270e-01 & 0 & 5.3422e+00 \\ 6.8793e-01 & -5.3422e+00 & 0 \end{bmatrix}$		
DMER =	[0 0 0]	-3.4396e+01	2.6711e+02 0]

Mode 3			
$C =$	$\begin{bmatrix} 0 & -7.6309e-01 & -2.5392e+00 \\ 7.6309e-01 & 0 & 9.1677e-01 \\ 2.5392e+00 & 9.1677e-01 & 0 \end{bmatrix}$		
DMER =	[0 0 0]	-1.2696e+02	-4.5838e+01 0]

X-/Ka-Band Dichroic Plate Design and Grating Lobe Study

J. C. Chen

Ground Antennas and Facilities Engineering Section

An X-/Ka-band dichroic plate is needed for simultaneously receiving X-band and Ka-band in the DSS-13 Beam Waveguide Antenna. The plate is transparent to the allocated Ka-band downlink (31.8–32.3 GHz) and the frequency band for the Mars Observer Ka-band Beacon Link Experiment (KABLE) (33.6–33.8 GHz), while at the same time reflecting the X-band downlink (8.4–8.5 GHz). The design is made using a computer program for dichroic plates with rectangular holes. The theoretical performance of the X-/Ka-band dichroic plate is presented. A study of the grating lobe problem is also included in this article.

I. Introduction

The bandwidth requirement for the X-/Ka-band dichroic plate at DSS 13 is 6.1 percent at Ka-band (31.8–33.8 GHz), whereas the desired bandwidth is 8.72 percent, which includes the Ka-band uplink (34.2–34.7 GHz). The holes on the dichroic plate have to be big enough for the low-frequency end of the Ka-band, but packed tight enough to prevent grating lobes at the high-frequency end. The wider the bandwidth, the more severe the grating lobe problem becomes. Also, there is a mechanical constraint for the wall thickness to ensure that the plate can support itself. There are over 10,000 holes for a 28-inch-diameter circular X-/Ka-band dichroic plate.

II. X-/Ka-band Dichroic Plate Design

The initial dimensions of an X-/Ka-band dichroic plate are scaled from the test S-/X-band dichroic plate, which is used to verify the computer code [1]. The reflection losses of the two orthogonal linear polarizations (TE and TM linear polarization) are not optimized in the test plate.

Therefore, the cell size, the hole size, and the thickness of this plate are adjusted to meet the requirement, which is ≤ 0.2 -dB reflection loss at the Ka-band Beacon Link Experiment (KABLE) frequency. The angle of incidence is $\theta = 30.0$ deg, $\phi = 0.0$ deg. The TE polarization is affected more by the changes of the X components of cell size and hole size than by the changes of Y components, while the TM polarization is affected more by the changes of the Y components of cell size and hole size than by the changes of X components for the angle of incidence $\phi = 0.0$ deg. The wall thickness is limited to a 0.008-in. minimum due to mechanical constraints.

The optimized design of the X-/Ka-band dichroic plate employs a rectangular hole of size 0.200 in. (H_x) by 0.203 in. (H_y); cell size 0.244 in. (D_x) by 0.211 in. (D_y) with a skew angle $\Omega = 60.0$ deg; and plate thickness (t) = 0.365 in. (Fig. 1). The amplitude and phase of the transmission and reflection coefficients are shown in Figs. 2–5. Reflection loss (transmission coefficient, dB) is less than 0.003 dB at the Ka-band downlink, 0.2 dB at the KABLE frequency, and 0.17 dB at the Ka-band uplink for both the

TE and TM linear polarizations. Since the received wave is circularly polarized, the reflection loss of the circular polarized wave at the KABLE frequency is around 0.1 dB, which meets the requirement. The phase shift between two linear polarizations is 3.4 deg to 4.0 deg at the Ka-band downlink, -0.2 deg to 0.5 deg at the KABLE frequency, and 1.8 deg to 3.8 deg at the Ka-band uplink (Fig. 6).

The tolerance of all the dimensions is 0.001 inch. Table 1 indicates the sensitivity of the reflection loss with respect to the hole size, which is the least controllable factor. The conductivity loss of the TE_{10} waveguide mode is 0.013 dB at the Ka-band downlink, 0.01 dB at the KABLE frequency, and 0.009 dB at the Ka-band uplink for a waveguide of size 0.203 in. by 0.200 in., thickness 0.365 in., and conductivity of copper 5.8×10^7 mhos/m [2].

III. Grating Lobe Study

Generally, the angle of transmission is equal to the angle of incidence (principal direction). But for certain circumstances, higher order modes are excited and some of the transmitted power leaves at an angle other than the principal direction. These excited higher order modes are called grating lobes and are similar to the grating lobes in phased array antennas. The conditions of the existence of a grating lobe depend on the angle of incidence, the frequency, and the size and skew angle of the cells on the plate. Since the incident wave is a horn pattern rather than a perfect plane wave, grating lobes may be generated by portions of the horn radiation that strike the plate at large angles of incidence.

The study of the grating lobe effects on a dichroic plate/feedhorn system is carried out in a series of steps. First, grating lobes that may be excited by a plane wave illumination are discussed. Next, expansion of a feedhorn pattern in terms of a group of plane waves traveling at different angles is described. Finally, the grating lobes that are excited by a horn pattern are determined.

A. Structure Dependence

For a certain angle of incidence ϕ , as θ is increased beyond some critical value ($\geq \theta_g$), grating lobes may be excited. Figures 7-9 show θ_g versus ϕ for different frequencies and cell sizes. The X-axis and Y-axis represent the angles ϕ and θ , respectively. For example, the point (0,55) in Fig. 7 indicates that when $\phi = 0.0$ deg, grating lobes are generated for $\theta \geq 55.0$ deg for the given cell size with respect to the wavelength. The θ_g curve changes when the skew angle of the cell changes (Fig. 7). Also, the higher the frequency, the smaller the θ_g for a given cell

size. Figure 8 shows the θ_g curve for wavelength $1.2 D$ to $1.6 D$ on a cell of $D_x = D$, $D_y = 0.866 D$, and $\Omega = 60$ deg, where D is a unit length. Figure 9 shows how the sizes of cells affect the value of θ_g .

B. Horn Pattern Expansion

The far-field horn pattern can be represented as a plane wave expansion, i.e., a group of plane waves traveling with different amplitudes at different angles. If the horn is placed at θ_0 , the different angles of incidence of these plane waves on the dichroic plate may be calculated through a coordinate transformation from the horn coordinate system to the dichroic-plate coordinate system. The far-field horn pattern is assumed to be symmetric in ϕ . Figures 10 and 11 show the angle of incidence on a dichroic plate from a horn with a symmetric pattern placed at $\theta_0 = 15.0$ deg, $\phi_0 = 0.0$ deg, and $\theta_0 = 30.0$ deg, $\phi_0 = 0.0$ deg, respectively. Each curve is for a constant θ_h in the horn coordinate system. For a 22-dB horn, $\theta_h = 12.0$ deg is about -12 dB with respect to the on-axis peak and $\theta_h = 14.0$ deg is about -14 dB.

C. Excitation of Grating Lobes by Horn Patterns

The grating lobe plot and the horn pattern plot need to be overlapped to determine the severity of the grating lobe problem for a given horn/dichroic plate geometry. For example, from the test S-/X-band dichroic plate experiment, the discrepancies between the calculation and measurement start above 8.6 GHz for TM polarization (Fig. 12) [1]. The conditions of the experiment are depicted in Fig. 13. The principal direction of the radiation from the 22-dB X-band horn strikes the plate at an angle of $\theta_0 = 30.0$ deg, $\phi_0 = 90.0$ deg. For 8.6 GHz, the θ_g curve intersects the $\theta_h = 14.0$ deg contour of the horn pattern, which corresponds to -14 dB with respect to the center radiation. Therefore, the disagreement between theory and experiment can be attributed to grating lobes excited by horn radiation that are at least 14 dB below the peak. Furthermore, Fig. 13 indicates that the grating lobe problem would be less severe if the principal direction of the horn was shifted to $\phi_0 = 0.0$ deg for $\theta_0 = 30.0$ deg.

For the X-/Ka-band dichroic plate, the horn is tilted 30 deg from normal at $\phi = 0.0$ deg (Fig. 14). The θ_g curve for 33.8 GHz (the high end of the KABLE frequency band) intersects with $\theta_h > 16.0$ deg curves, which is equivalent to 16 dB below the radiation peak, while the θ_g curve for 34.7 GHz (the high end of the Ka-band uplink) intersects partially with the $\theta_h = 14.0$ deg contour. This implies that there may be grating lobe effects at the high end of the Ka-band uplink. Since there is no noise consideration in the uplink band and only a slight power increase would be

required to compensate for the additional loss, this grating lobe problem is not considered to be severe.

IV. Conclusion

The design of an X-/Ka-band dichroic plate is presented. The theoretical calculations show excellent per-

formance at the Ka-band downlink frequency and show a reflection loss that meets the requirement at the KABLE frequency. The analysis of the grating lobe effects predicts that there is no grating lobe problem at either the Ka-band downlink or KABLE frequencies, yet there may be grating lobe effects at the high end of the Ka-band uplink. This needs to be verified by experiments.

References

- [1] J. C. Chen, "Analysis of a Thick Dichroic Plate with Rectangular Holes at Arbitrary Angle of Incidence," *TDA Progress Report 42-104*, vol. October-December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 9-16, February 15, 1991.
- [2] R. E. Collin, *Foundations for Microwave Engineering*, New York: McGraw-Hill Physical and Quantum Electronics Series, p. 103, 1966.

Table 1. Tolerance study of hole size

Frequency, GHz	Polarization	Hole size, in.				
		$H_x = 0.200$ $H_y = 0.203$	$\Delta H_x = -0.001$	$\Delta H_x = +0.001$	$\Delta H_y = -0.001$	$\Delta H_y = +0.001$
		Reflection loss, dB	Δ Reflection loss, dB	Δ Reflection loss, dB	Δ Reflection loss, dB	Δ Reflection loss, dB
32.3 (Ka-band downlink)	TE	0.023	+0.002	+0.016	+0.003	-0.002
	TM	0.025	-0.001	+0.002	-0.015	+0.019
33.8 (KABLE)	TE	0.007	+0.003	-0.004	+0.000	-0.001
	TM	0.199	-0.005	+0.004	+0.001	-0.003
34.7 (Ka-band uplink)	TE	0.078	-0.017	+0.019	+0.002	-0.002
	TM	0.096	-0.002	+0.003	+0.009	-0.008

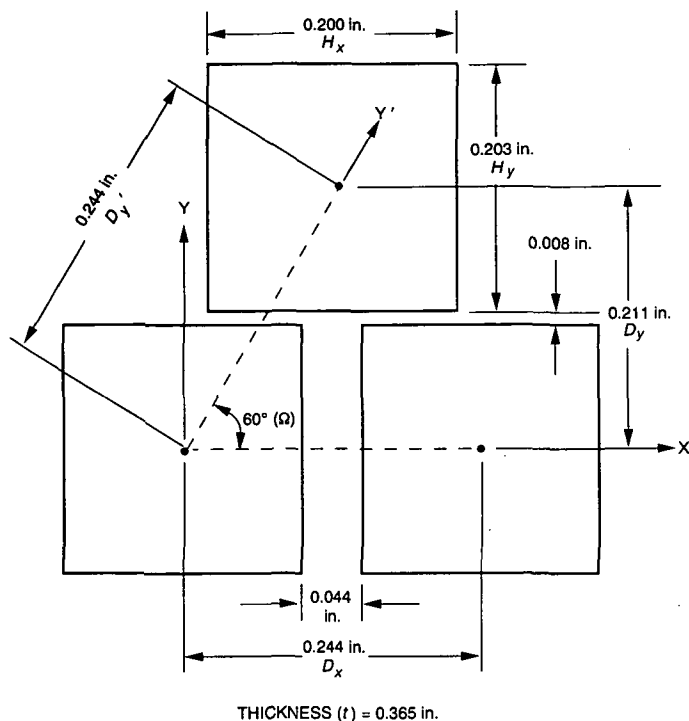


Fig. 1. Geometry of X-/Ka-band dichroic plate with rectangular holes.

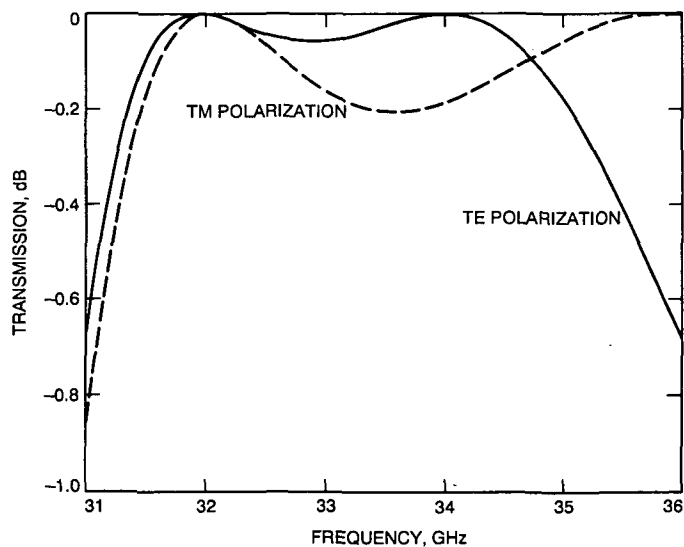


Fig. 2. Theoretical transmission versus frequency for X-/Ka-band dichroic plate for TE and TM polarization.

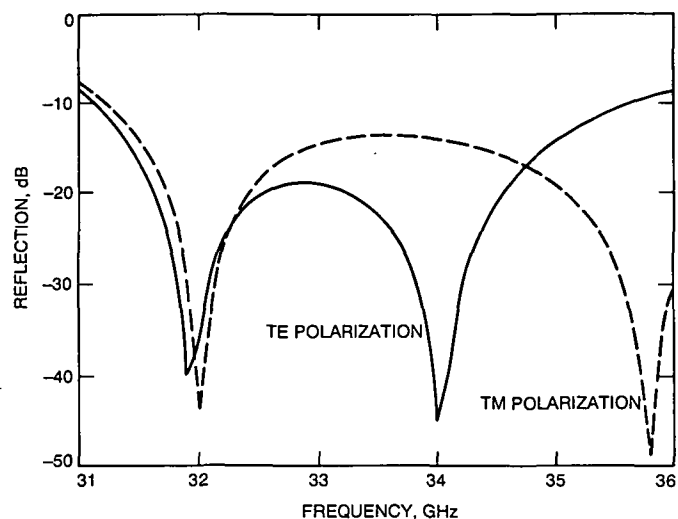


Fig. 3. Theoretical reflection versus frequency for X-/Ka-band dichroic plate for TE and TM polarization.

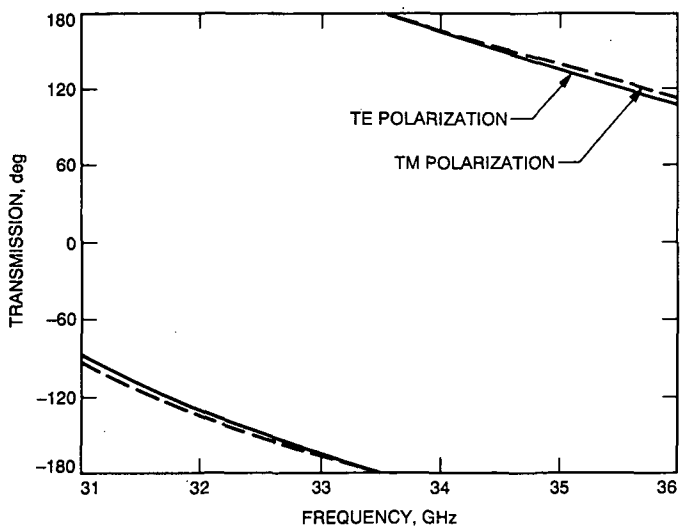


Fig. 4. Theoretical transmission phase versus frequency for X-/Ka-band dichroic plate for TE and TM polarization.

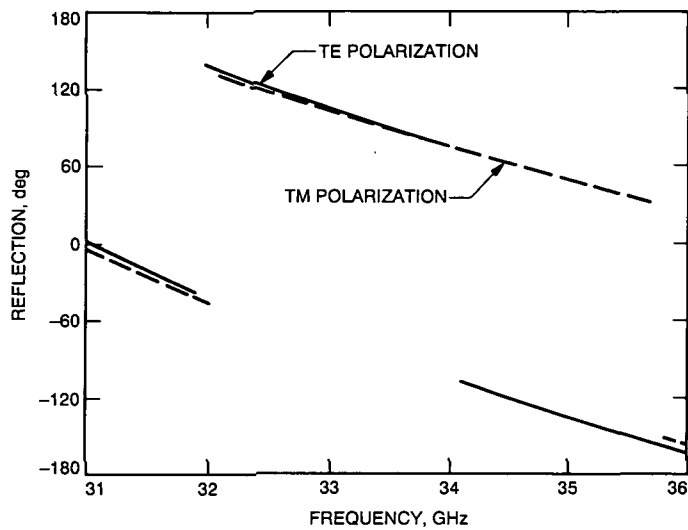


Fig. 5. Theoretical reflection phase versus frequency for X-/Ka-band dichroic plate for TE and TM polarization.

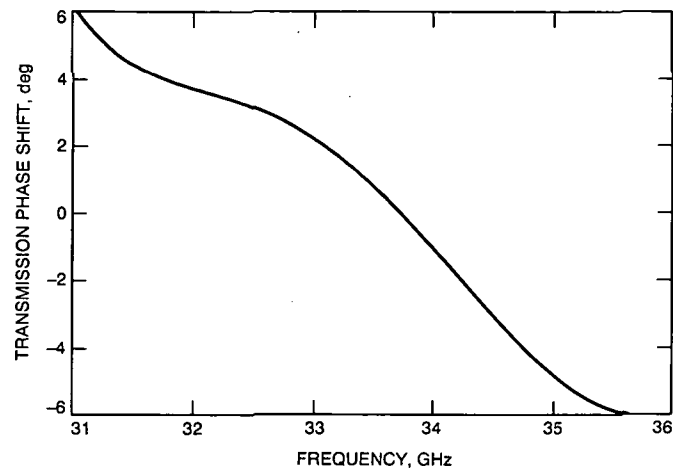


Fig. 6. Phase shift between TE and TM polarization versus frequency for X-/Ka-band dichroic plate.

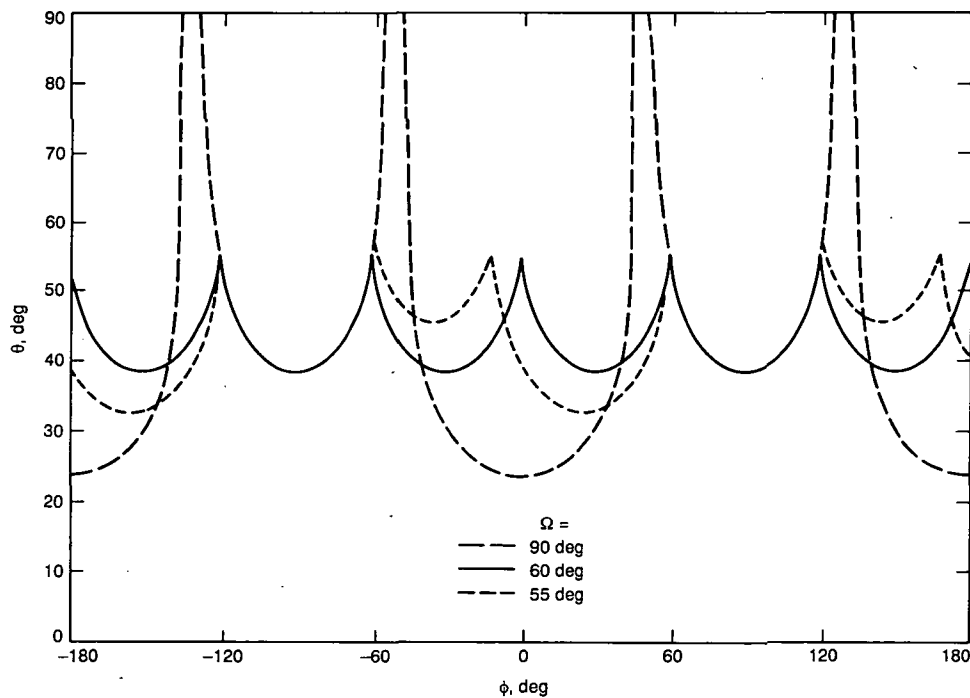


Fig. 7. Grating lobe curves for various skew angles (Ω) at a certain frequency (wavelength = $1.4 D$) with cells of $D_x = D$, $D_y = 0.866D$, where D is a unit length.

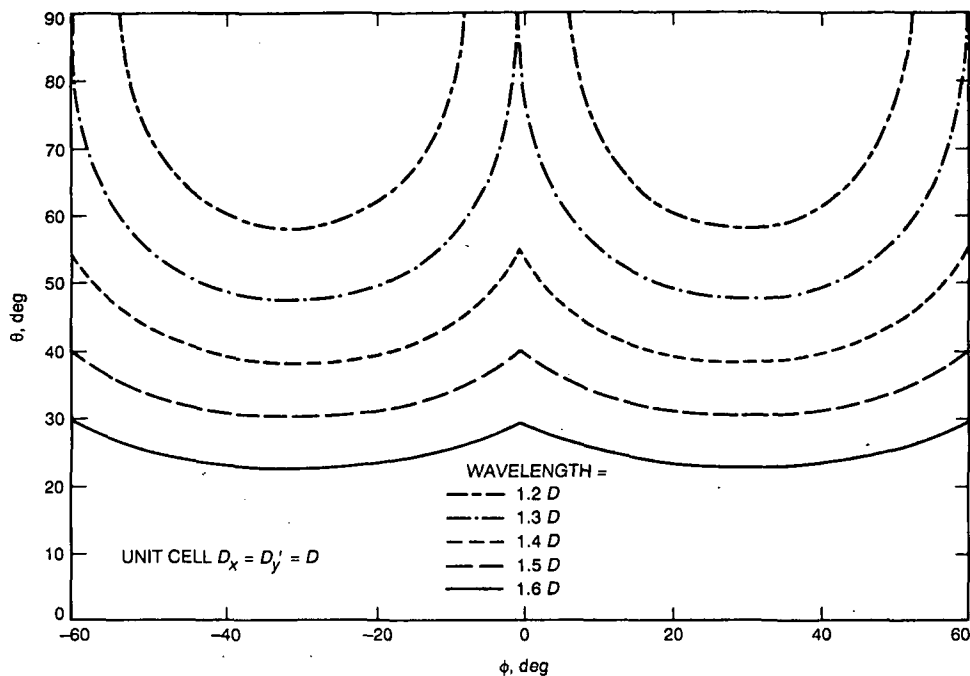


Fig. 8. Grating lobe curves for various frequencies (wavelength = $1.2 D$ to $1.6 D$) with a cell of $D_x = D$, $D_y = 0.866D$, and $\Omega = 60.0$ deg, where D is a unit length.

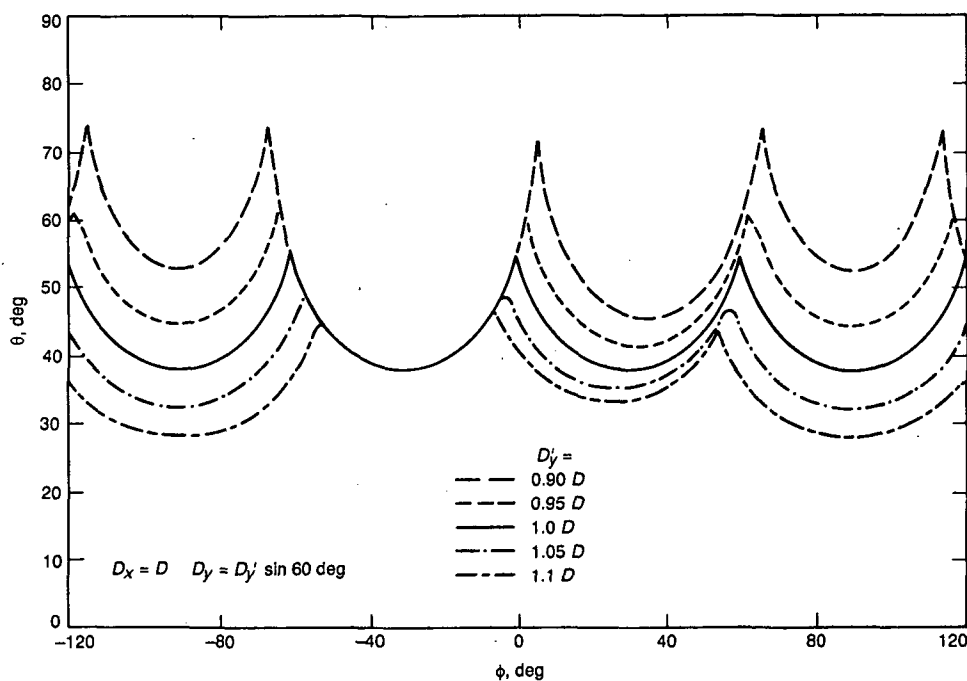


Fig. 9. Grating lobe curves with different Y' components for a cell of $D_x = D$, $D_y' = D$, and $\Omega = 60.0$ deg, where D is a unit length.

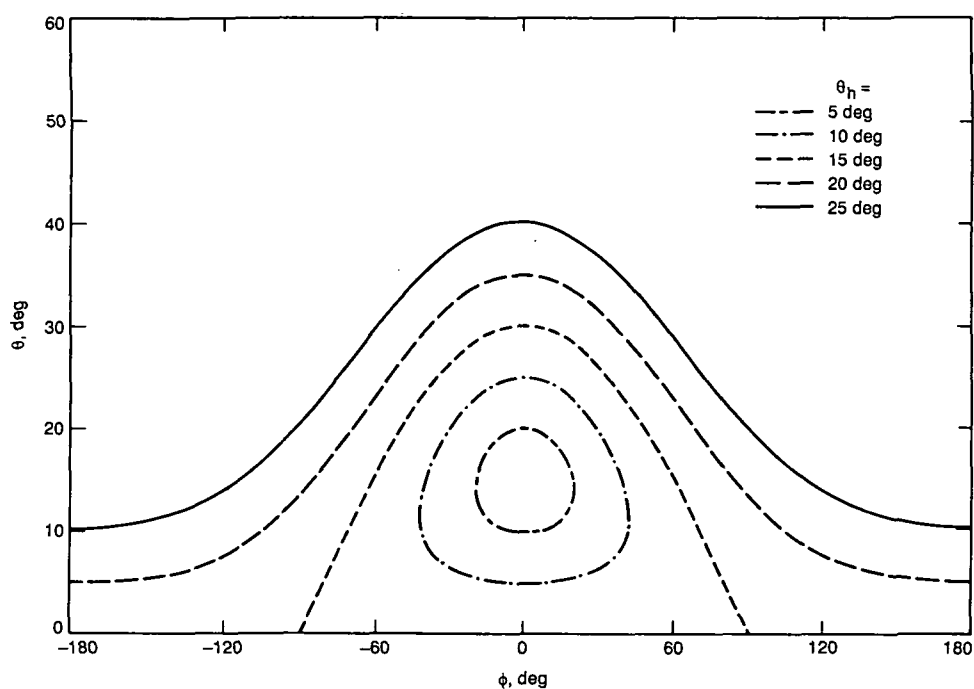


Fig. 10. Contour plot of horn pattern in terms of angles of incidence for a horn with principal angle at $\theta_0 = 15.0 \text{ deg}$, $\phi_0 = 0.0 \text{ deg}$.

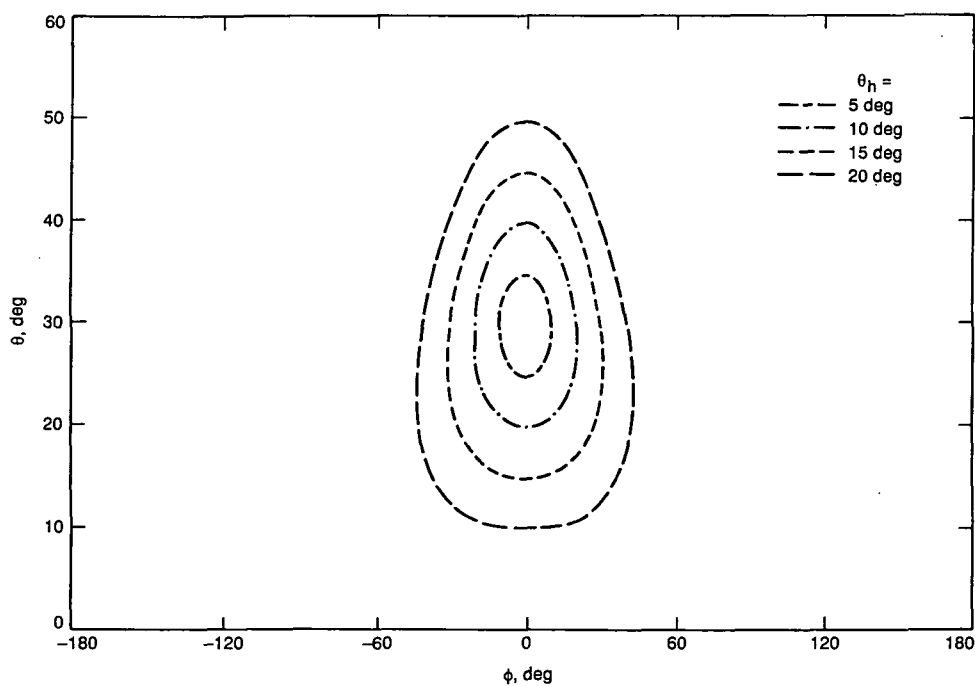


Fig. 11. Contour plot of horn pattern in terms of angles of incidence for a horn with principal angle at $\theta_0 = 30.0 \text{ deg}$, $\phi_0 = 0.0 \text{ deg}$.

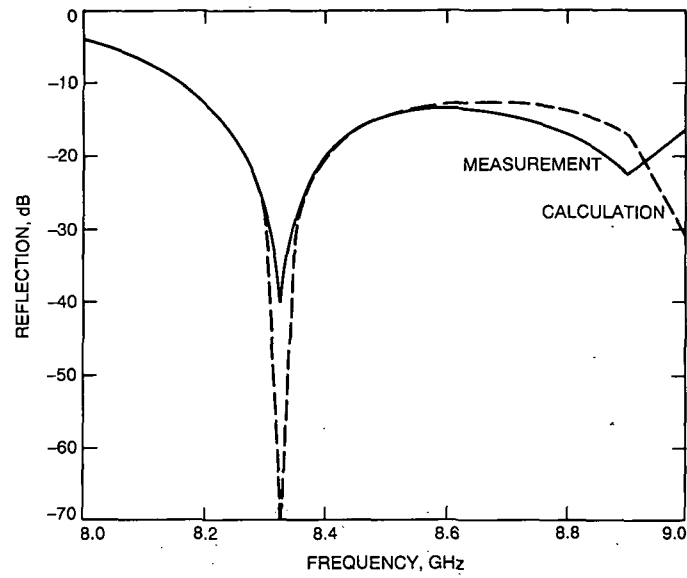


Fig. 12. Measured and calculated reflection versus frequency for the test S-/X-band dichroic plate.

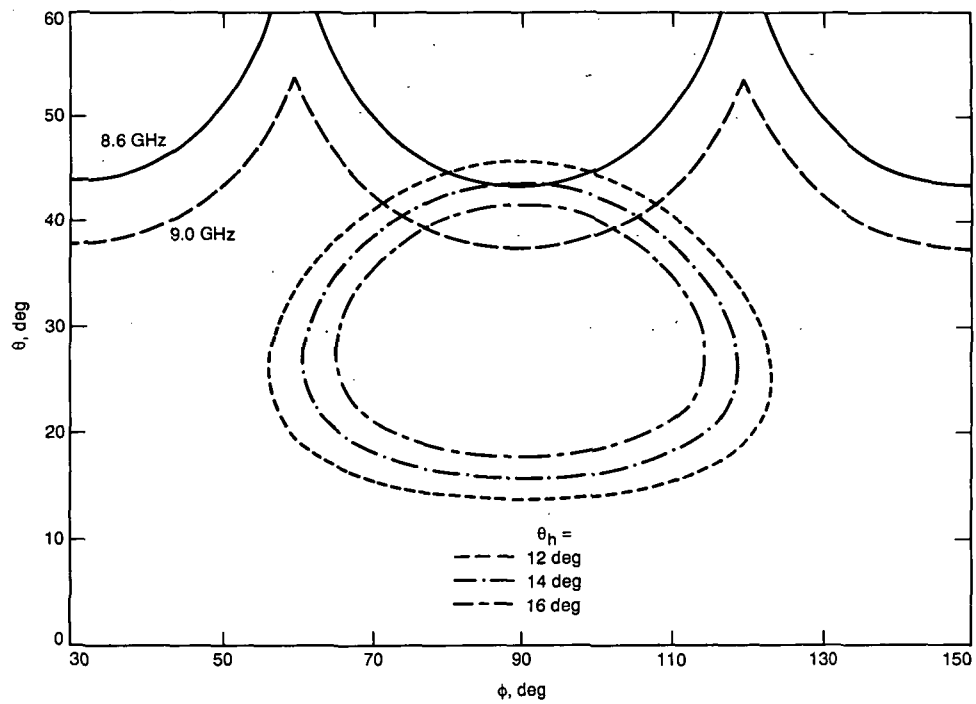


Fig. 13. Excitation of grating lobes by the feedhorn pattern for the test S-/X-band dichroic plate.

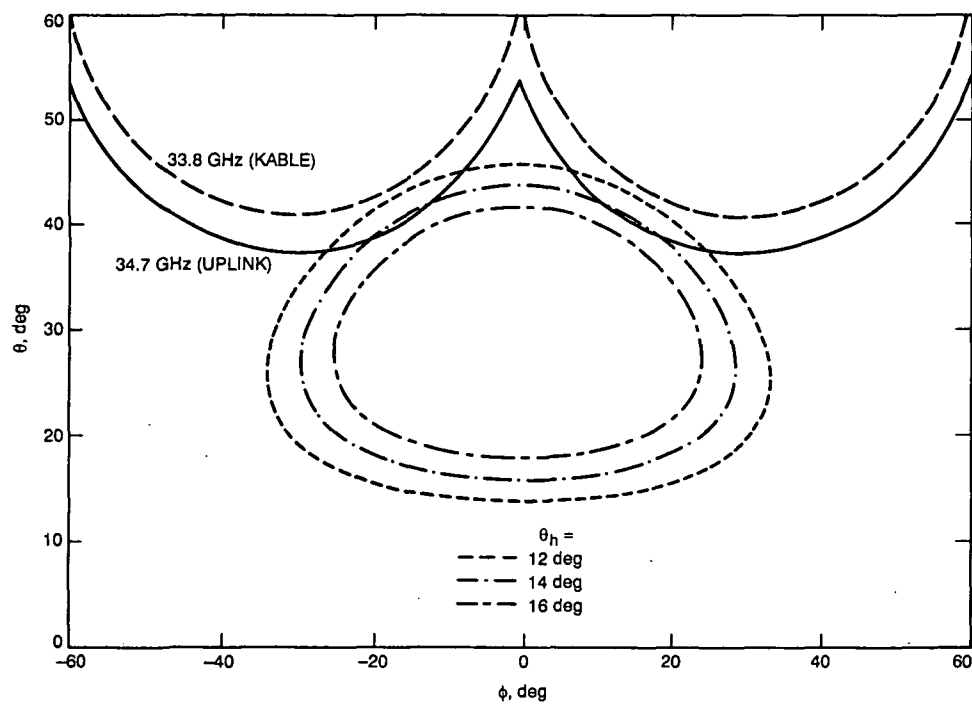


Fig. 14. Excitation of grating lobes by the feedhorn pattern for the X-/Ka-band dichroic plate.

1994025714

N94-29617

TDA Progress Report 42-105

442586
May 15, 1991

Site Comparison for Optical Visibility Statistics in Southern California

K. Cowles

Communications Systems Research Section

Negotiations are under way to locate an atmospheric visibility monitoring (AVM) observatory at Mount Lemmon, just north of Tucson, Arizona. Two more observatories will be located in the southwestern U.S. The observatories are being employed to improve a weather model for deep-space-to-ground optical communications. This article explains the factors considered in choosing a location and recommends Table Mountain Observatory as the location for another AVM facility.

I. Introduction

Ground-based receivers of optical communications signals from planetary spacecraft must contend with the adverse effects of the Earth's atmosphere. Spatially diverse receivers will improve link performance by providing at least one site with favorable atmospheric conditions.

The atmospheric visibility monitoring (AVM) project is designed to create an improved weather model for cloud-cover statistics over the southwestern United States. An improved model will help predict the performance of ground-based optical receivers projected for construction. A preliminary model already predicts that it is possible to have clear skies at at least one of three sites 95 percent of the time if the sites are properly distributed throughout the region [1]. In order to determine a more exact weather model, the observatories are being located in areas with a low correlation of cloudy skies, and where facilities already exist to supply roads, power and telephone lines. Several of the candidate sites are already locations of astronomical observatories.

Southern California is known to have favorable weather patterns that differ from those in Arizona and states farther east. This article describes the characteristics of, and recommendations for, a California site selection. Sites considered were Goldstone, Table Mountain Observatory (TMO), Mount Wilson, and Mount Laguna. Their locations are shown in Figure 1.

II. Site Characteristics

Five major site factors were considered, and each given a weighting relative to its importance to AVM. The factors are listed in Table 1. Candidate sites were then rated on a 1-10 scale for each factor. The ratings were in turn weighted and summed to give an overall rating for the site [2].

A. Probability of Clear Skies

Two factors are being considered within this characteristic. First, a site should exhibit a high annual percentage

of clear skies. Second, it should exhibit a low correlation of cloudy skies with the Mount Lemmon site. Annual percentage of clear skies can be determined from ground and satellite data that have been compiled over several years by organizations such as the National Weather Service, the National Climatic Data Center, and the Solar Energy Research Institute. These sources have produced the information in Table 2.

The data in the following paragraphs were compiled by Don Wylie of the University of Wisconsin from three years of Geostationary Operational Environmental Satellite/Visible and Infrared Spin Scan Radiometer Atmospheric Sounder (GOES/VAS) data.¹ The resolution of his data is limited to a 2-deg \times 2-deg box; consequently, TMO and Mount Wilson fall within the same box. The data are for January and July only; however, they tend to give a sufficient indication of what summer and winter conditions are like.

In January, if Mount Lemmon was cloudy, Goldstone was clear 59 percent of the time, TMO and Mount Wilson were clear 52 percent of the time, and Mount Laguna was clear 23 percent of the time. If Goldstone was cloudy, Mount Lemmon was clear 54 percent of the time. If TMO or Mount Wilson was cloudy, Mount Lemmon was clear 63 percent of the time. If Mount Laguna was cloudy, Mount Lemmon was clear 44 percent of the time.

In July, if Mount Lemmon was cloudy, Goldstone was clear 87 percent of the time, TMO and Mount Wilson were clear 84 percent of the time, and Mount Laguna was clear 67 percent of the time. If Goldstone was cloudy, Mount Lemmon was clear 50 percent of the time. If TMO or Mount Wilson was cloudy, Mount Lemmon was clear 50 percent of the time. If Mount Laguna was cloudy, Mount Lemmon was clear 21 percent of the time.

These data indicate that Goldstone, TMO, and Mount Wilson exhibit a much lower correlation of cloudy skies with Mount Lemmon than Mount Laguna does. Other data from Wylie indicating the probability of clear skies, cirrus clouds, and opaque clouds for these sites in January and July are shown in Table 3. These data show that Goldstone has a greater probability of clear skies than any of the other sites. In fact, the number of clear days at Goldstone is comparable to the number of clear days at Mauna Kea, Hawaii, a site well known for its clear skies [5]. Mount Laguna has a higher probability of clear skies than TMO and Mount Wilson; however, the correlation

data show that their cloudiness is more correlated with cloudiness at Mount Lemmon.

Based on this information, the following ratings were given for probability of clear skies: Goldstone: 10, TMO: 9, Mount Wilson: 9, and Mount Laguna: 6.

B. Low Particulate Scattering

Particulate scattering arises from particles in the atmosphere. Scattering will attenuate light transmission and increase daytime sky radiance. Sites at higher elevations tend to exhibit less scattering because the atmosphere is thinner at higher altitudes and there is a shorter path from the ground to space. Large measures of precipitable water vapor (PWV) are an indication of more scattering and higher atmospheric extinction. Measurements of PWV are available from sites involved in infrared astronomy. Sites with vegetation tend to have less dust blown up by winds. Smog, fog, and haze also contribute to higher particulate scattering. LOWTRAN7, a low-resolution propagation model for calculating atmospheric transmittance and background radiance [6], was used to help quantify site factors that were examined.

Goldstone is at an elevation of 1040 m, much lower than any of the other potential California sites. Since it is in a desert area, vegetation is minimal. Precipitable water vapor measurements are not available for Goldstone. LOWTRAN7 calculations were done for all of the sites to predict both transmission and background sky radiance (Figs. 2-8). Maximum transmission at Goldstone was found to be 88 percent at zenith at 1.064 μm . The graphs indicate that the three other sites had a much higher transmission at all wavelengths and zenith angles. A tropospheric-haze model with 50-km visibility was used for the mountain locations, whereas a rural-haze model with 23-km visibility was used for the Goldstone calculations. These models were used because Goldstone lies in an area level with most of its surroundings, whereas the mountain sites will be above inversion layers that keep down the valley haze and increase visibility.

Daytime background sky radiance increases with increased particulate scattering. More scattering not only decreases signal transmission, but also increases background noise during the day. LOWTRAN7 calculations for radiance overall were highest at 0.532 μm and lowest at 1.064 μm , and radiance at Goldstone was approximately twice as high as that for the other sites. The coordinate system for radiance calculations is shown in Fig. 5. Figures 6-8 show the worst-case calculations, with the Sun at 30 deg from zenith (the highest it gets in the sky at that

¹ D. Wylie, personal communication, University of Wisconsin, Madison, Wisconsin, March 27, 1990.

latitude), and looking 30 deg from the Sun (the closest the AVM telescopes will be pointed to the Sun). Worst-case radiance in $W/(cm^2 \cdot sr \cdot \mu m)$ at $0.532 \mu m$ for Goldstone is 0.025, for TMO is 0.012, for Mount Wilson is 0.013, and for Mount Laguna is 0.013.

An experiment was done at TMO and Goldstone to test the performance of a charge-coupled device (CCD) camera and photometer when distinguishing stars from sky background during the day. The telescope was pointed at the star Altair, about 75 deg from the Sun, much farther than the worst-case calculations (30 deg). However, the differences in radiance could still be seen. Figure 9 is a graph of photometer readings of the background sky near Altair as it moved across the sky on the mornings of April 11 and 12, 1990. The differences in the number of photons collected are caused by the greater amount of background light at Goldstone. Increased haze resulting from the lower elevation and desert environment increase the amount of scattered sunlight, which in turn creates the background light. Readings for Goldstone averaged 517 counts higher than readings for TMO, a significant amount considering the total readings ranged from 1300 to 3000.

Goldstone's location on the desert floor inhibits detection of stars against their background because haze settles in valleys and on plains. Variations in transmission at all three wavelengths are shown in Fig. 10. The signal transmission is lower and daytime sky background noise is higher when compared with sites located above an inversion layer that confines the haze to lower altitudes.

TMO is at 2290 m, the highest of the California sites. Precipitable water vapor measurements indicate an average rainfall of from 2 to 8 mm annually [7]. The peak is covered with tall pine trees and low brush. It is above an inversion layer that confines valley smog, fog, and haze. LOWTRAN7 calculations indicate successful light transmission comparable to that of the other mountain sites, the maximum being 95 percent at zenith for $1.064 \mu m$. Figure 11 compares all three wavelengths, showing how transmission varies with wavelength and zenith angle.

Mount Wilson is at an elevation of 1740 m. Despite its proximity to Los Angeles, Mount Wilson does not have a significant problem with smog, since an inversion layer usually exists well below the summit. Cloud cover and coastal fog measurements taken for Los Angeles are not good indicators of conditions on Mount Wilson, since it is often above the low clouds and fog that cling to the coastline. The vegetation is similar to that at TMO. LOWTRAN7 light transmission and radiance calculations are similar to those for TMO. Precipitable water vapor measurements are not available.

Mount Laguna is at an elevation of 1860 m. Annual precipitable water vapor measurements average 5.2 mm [8]. LOWTRAN7 transmission calculations are similar to those for TMO and Mount Wilson.

Based on the above information, the following ratings were given to the sites for low particulate scattering: Goldstone: 4, TMO: 9, Mount Wilson: 8, and Mount Laguna: 9.

C. Suitability for a Future Large Optical Reception Station

The purpose of the AVM project is to enhance the cloud-cover model for optical communications from deep space to ground. Data will be collected from the AVM sites on atmospheric transmission and cloud cover. Should a site prove to be a favorable location for receiving communications from spacecraft, it will be a good candidate location for a receiving station. Current JPL projects include the design of a 10-meter segmented optical receiver [9]. Therefore, if data are to be taken on cloud-cover conditions at a specific site, it is desirable to choose one that can eventually accommodate a large receiver.

Goldstone is an excellent location for a large receiving station because it is already the site of the current U.S. node of the Deep Space Network (DSN) and can quite easily accommodate another receiver.

TMO is another JPL site that can accommodate a larger receiver. Dan Sidwell, the JPL site manager, has explained that the most desirable site on the mountain for a large receiver will be available in a few years should JPL decide to locate one at TMO. The large number of trees in the area would not significantly block observation because the view to the south, and hence of the ecliptic plane, is not obstructed. Difficulties in locating a larger receiver at TMO would arise from limited space available for buildings other than the AVM telescope itself. Lights from Mountain High, a nearby ski area, do not create problems for the AVM systems because the AVMs are designed to operate during the day. Any light creating difficulties for a large receiver can be filtered out since the ski area uses mercury vapor lamps that emit light at only a few narrow bands.

Mount Wilson has the advantage of being very close to JPL. In addition, it is not overcrowded with telescopes and has enough space for a large AVM facility. A disadvantage is that it is not already a JPL facility, as are Goldstone and TMO, so more negotiations would be involved in setting up a larger facility there.

Mount Laguna is currently operated by San Diego State University (SDSU) in the Cleveland National Forest. The ridge where the SDSU observatory is located still has abundant room for additional telescopes. But Mount Laguna often receives weekend visitors from nearby campgrounds because the observatory allows them to view the skies through a small antique telescope. There might be a conflict of interest in having a large number of people nearby when a large transmitter or receiver is being operated. Dr. R. Angione, director of Mount Laguna Observatory, does not foresee any problems with locating a large receiver there, however.

One problem common to all of the candidate sites is that of gaining approval from the Federal Aviation Administration (FAA) to restrict air space to transmit optical signals. This is because all of the sites are located near already restricted military air space (Goldstone), commercial air routes (TMO and Mount Wilson), or unrestricted space frequented by military and private aircraft (TMO and Mount Laguna). Transmitting at low angles to the horizon could result in laser penetration into these areas, which would interfere with air traffic. Solving this problem is beyond the scope of the AVM project, but air-space availability must be a factor in selecting a site for a large receiver.

Based on the above information, the following ratings were given to the sites for suitability of a future large optical reception station: Goldstone: 10, TMO: 9, Mount Wilson: 8, and Mount Laguna: 8.

D. Availability of Security and Maintenance

The availability of security and maintenance at a site are also considered because JPL personnel will not be able to frequently monitor the AVM observatory. Observatories must be protected from vandalism as well as accidents caused by curious hikers (therefore, it would be a good idea to put a fence around the AVM buildings, no matter where they are located). Minimal maintenance, such as snow removal and checks for weather damage or mechanical failure, is also necessary. It would also be useful to have someone available to call on in case a problem is detected from JPL. The designated individual could check the site and possibly close the roof until someone arrived to fix the problem. How often periodic checks would be required has not yet been determined.

Security at Goldstone is not a problem. Unauthorized personnel would not normally be near the observatory. But maintenance might be somewhat difficult because there is currently no one assigned to make routine

checks of facilities. Nonetheless, if the observatory is located near the Deep Space Station (DSS) 13 (Venus site), maintenance personnel will be close enough to add a routine visit to their daily tasks. A better site in terms of atmospheric conditions would be above the Apollo station, on a mesa about 200 feet higher than the surrounding area; however, maintenance personnel would be less available to make checks at that site.

TMO does not have the advantage of being restricted to JPL personnel. However, security has not been a significant problem there. The facility is staffed 24 hours a day, 7 days a week. Maintenance at TMO would be easily handled by the staff members who traverse the site every two hours on a fire-watch patrol.

Mount Wilson is not a very secure site since it quite frequently has visitors from the Los Angeles area. Therefore, a protective fence definitely would have to be erected around an AVM observatory. Another problem is that the only personnel who are available to make routine maintenance checks are observing technicians, who have many other duties. If a problem occurs, JPL personnel could drive to Mount Wilson since it is so close. However, this would take some time and would not be practical on a daily basis.

Mount Laguna Observatory is not secured from the rest of the National Forest by a fence. As mentioned before, campers are often present. Dr. Angione reported that only three minor events of vandalism have occurred in the past twenty years, none of which harmed any of the telescopes or other major equipment. Still if an AVM observatory is located at Mount Laguna, it would be a very good idea to have a secure fence around it for both safety and security reasons. The site superintendent could maintain the AVM observatory Monday through Friday when he is working. There would be no one on the site during weekends; however, the superintendent lives a mile from the observatory, so he might be available for emergencies.² Maintenance would be mostly limited to mechanical problems.

Based on the above information, the following ratings were assigned to the sites for availability of security and maintenance: Goldstone: 8, TMO: 9, Mount Wilson: 6, and Mount Laguna: 7.

E. Low Turbulence

Turbulence decreases the ability to obtain a focused image. Sites that exhibit low turbulence are usually at the

² R. Angione, personal communication, SDSU, San Diego, California, May 23, 1990.

top of a steep slope rising out of a valley in the direction of prevailing winds. Such sites also have no peaks in that direction higher than 1 km for at least 30 km. As with particulate scattering, the higher the elevation of the telescope, the better, because there is less atmosphere to blur images. Turbulence will be worse during daylight hours at any site due to the heat rising from the ground as it is warmed by the Sun. Turbulence is indicated by a site's capability of "seeing," a measurement of the blurring of an image.

Goldstone experiences greater turbulence than the other sites since it is at a low elevation in a desert, where the ground gets substantially warmer during the day, rather than at a high elevation with more vegetation. No good records of seeing have been kept at Goldstone since little optical work has been done there. However, a few amateur groups in the area have made some estimates of nighttime seeing. During the first part of the night, when the ground is cooling, seeing is generally poor, but it improves after midnight and then is generally between 2 and 10 arcsec. No measurements have been made during the day; however, seeing probably degrades greatly, due to the heat rising from the desert floor and the constancy of elevation in the area (if the site were 100 m or more above the surrounding desert, it would not be subject to some of the ground effects produced when looking at large zenith angles).

Seeing measurements have not been made regularly at TMO other than at its 24-inch telescope. There is some debate about how much of the seeing at this telescope is "dome-effect" seeing (light deflection thermally induced by air-temperature inhomogeneity in the observing path inside the dome). Measurements vary from 2 to 10 arcsec inside the 24-inch telescope dome. Some measurements were made using a 10-inch Celestron telescope both outside and inside the dome, to determine the variation in seeing. Measurements inside the dome were about three times worse than outside. This would indicate that seeing caused by the site, not the dome, is usually from less than 1 arcsec to 3 arcsec.

Seeing at TMO is not as good as at Mount Wilson because air does not flow to the mountain directly from the ocean. However, it is still good because the peak's elevation is above most ground features, and provides less atmosphere to blur images.

Mount Wilson exhibits the least turbulence. The laminar air flow from the ocean provides very nonturbulent conditions since no mountains obstruct it before it reaches Mount Wilson. Turbulent air occurs only when wind from the north crosses the San Gabriel mountains. Turbulence measurements indicate seeing to average 0.5 arcsec.

Mount Laguna does not have good records for average seeing measurements; however, Dr. Angione estimates that about 60 percent of all nights are photometric and that seeing is typically between 1 and 2 arcsec at those times. Occasionally, it has been better than 1 arcsec. The ridge on which the observatory stands runs north-south and is higher than any others to the west, so it exhibits laminar air flow from the ocean similar to that of Mount Wilson.³

Based on the above information, the following ratings were given to the sites for low turbulence: Goldstone: 6, TMO: 8, Mount Wilson: 10, and Mount Laguna: 8.

III. Rating

The ratings are summarized in Table 4.

IV. Conclusions

Table Mountain Observatory is the best location for the California AVM observatory. The excellent astronomical conditions at the site, combined with the convenience of its already being a JPL facility staffed full time and available for future use, make it the prime candidate for placement of the second observatory. The correlation of cloudy skies is also quite low between TMO and Mount Lemmon, Arizona.

No definite cost (lease, construction, maintenance, operations, etc.) information is available for any of the sites, therefore, cost was not a factor in this selection process. However, it seems reasonable to assume that locating a facility at TMO will be less costly than at a location where site rental fees would need to be paid, as at Mount Lemmon.

³ Ibid.

References

- [1] K. Shaik, "A Preliminary Weather Model for Optical Communications Through the Atmosphere," *TDA Progress Report 42-95*, vol. July-September 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 212-218, November 15, 1989.
- [2] K. Cowles, "Site Selection Criteria for the Optical Atmospheric Visibility Monitoring Telescopes," *TDA Progress Report 42-97*, vol. January-March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 235-239, May 15, 1989.
- [3] *Climatic Atlas of the United States*, Washington, D.C.: U.S. Department of Commerce, 1968.
- [4] *Solar Radiation Energy Resource Atlas of the United States*, Golden, Colorado: Solar Energy Research Institute, 1981.
- [5] P.-Y. Bely, "Weather and Seeing on Mauna Kea," *Pub. Ast. Soc. Pac.*, vol. 99, pp. 560-570, June 1987.
- [6] F.-X. Kneizys, E. P. Shettle, G. P. Anderson, L. W. Abreu, J. H. Chetwynd, J. E. Selby, S. A. Clough, W. O. Gallery, LOWTRAN7 computer code, Air Force Geophysics Laboratory, Hanscom Air Force Base, Massachusetts, 1989.
- [7] R. G. Roosen and R. Angione, "Variations in Atmospheric Water Vapor: Baseline Results from Smithsonian Observations," *Pub. Ast. Soc. Pac.*, vol. 89, pp. 814-822, December 1977.
- [8] R. Angione, "Precipitable Water Vapor at Mount Laguna Observatory," *Pub. Ast. Soc. Pac.*, vol. 99, pp. 895-898, August 1987.
- [9] K. S. Shaik and E. L. Kerr, "A Ten-Meter Optical Telescope for Deep Space Optical Communications," *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, vol. 1236, ed. L. D. Barr, Tuscon, Arizona, pp. 347-350, February 12-16, 1990.

Table 1. Major site factors

Factor	Percent
Low probability of cloud cover, fog, smog, and haze	30
Low particle scattering	20
Suitability of the site for a potential optical reception station	20
Low turbulence	15
Availability of security and maintenance	15
	100

Table 2. Annual clear-sky statistics

Factor	Candidate site		
	Goldstone	TMO/ Mt. Wilson	Mt. Laguna
Approximate annual sunshine [3], mean total hours	3600	3300	3400
Annual average daily direct normal solar radiation [4], MJ/m ²	24	20	21
Annual average daily opaque sky cover [4], tenths of sky	2.5	4	3.2

Table 3. Probability of clear skies, cirrus clouds, and opaque clouds

	January, percent			July, percent		
	Clear	Cirrus	Opaque	Clear	Cirrus	Opaque
Goldstone	49	20	31	67	13	20
TMO/Mt. Wilson	38	22	40	36	6	58
Mt. Laguna	33	21	46	46	9	45
Mt. Lemmon	39	23	38	26	33	42

Table 4. Rating the candidate sites

Factor	Weight, percent	Candidate site			
		Goldstone	TMO	Mt. Wilson	Mt. Laguna
Clear skies	30	10	9	9	6
Low scattering	20	4	9	8	9
Future suitability	20	10	9	8	8
Security and maintenance	15	8	9	6	7
Low turbulence	15	6	8	10	8
Total rating		7.9	8.85	8.30	7.45



Fig. 1. Sites investigated for locating atmospheric-visibility monitoring observatories.

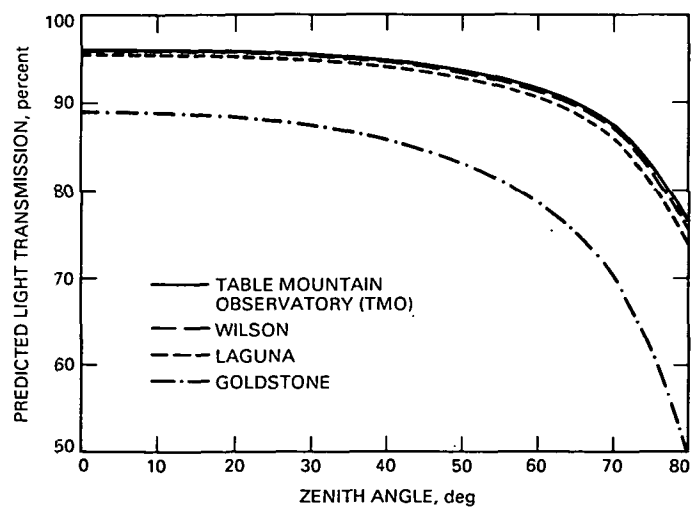


Fig. 2. Calculated light transmission at $1.064 \mu\text{m}$ for an average summer at candidate sites for atmospheric-visibility monitoring observatories.

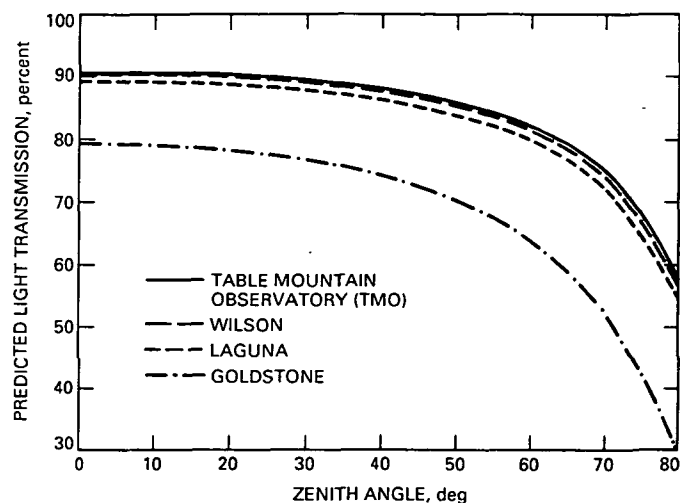


Fig. 3. Calculated light transmission at $0.830 \mu\text{m}$ for an average summer at candidate sites for atmospheric-visibility monitoring observatories.

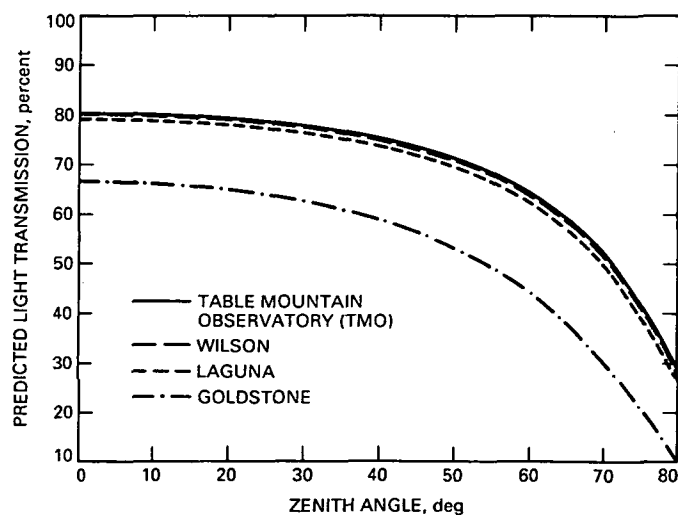


Fig. 4. Calculated light transmission at $0.532 \mu\text{m}$ for an average summer at candidate sites for atmospheric-visibility monitoring observatories.

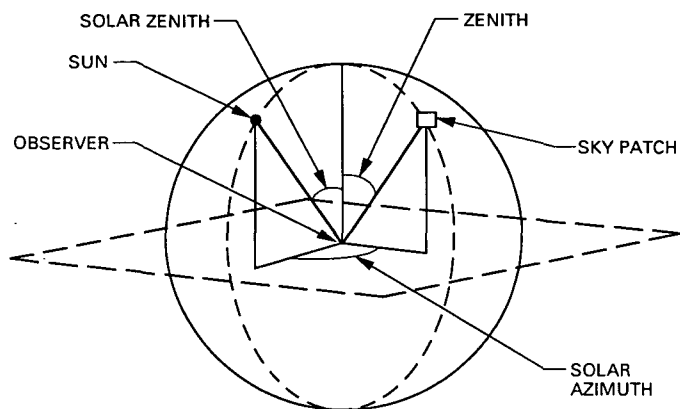


Fig. 5. Coordinate system for LOWTRAN7 radiance calculations.

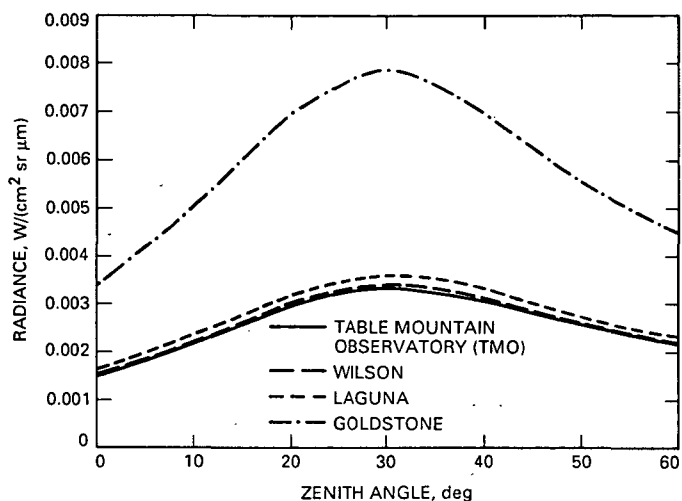


Fig. 7. Calculated radiance at 1.830 μm , in the presence of cirrus clouds with solar azimuth at 30 deg and solar zenith at 30 deg, for an average summer at candidate sites for atmospheric-visibility monitoring observatories.

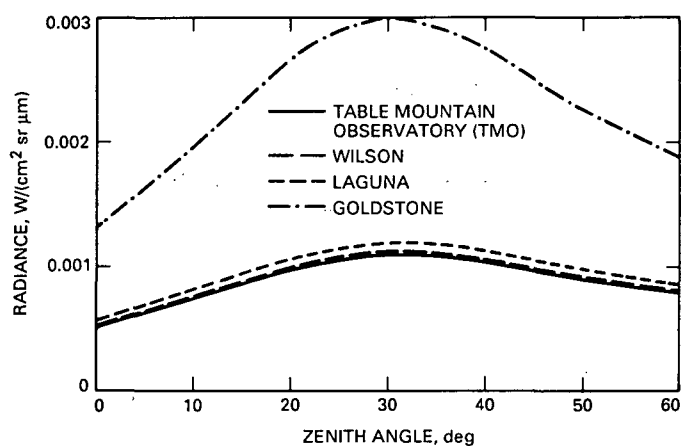


Fig. 6. Calculated radiance at 1.064 μm , in the presence of cirrus clouds with solar azimuth at 30 deg and solar zenith at 30 deg, for an average summer at candidate sites for atmospheric-visibility monitoring observatories.

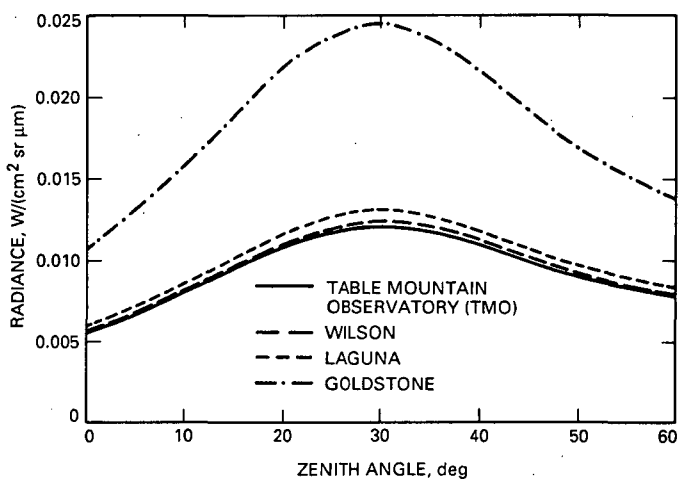


Fig. 8. Calculated radiance at 0.532 μm , in the presence of cirrus clouds with solar azimuth at 30 deg and solar zenith at 30 deg, for an average summer at candidate sites for atmospheric-visibility monitoring observatories.

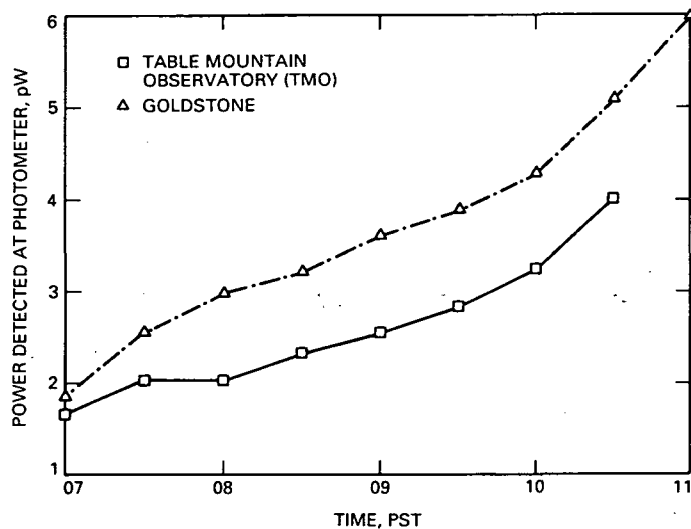


Fig. 9. Sky background photometer readings at 1.064 μm for Goldstone and Table Mountain on April 12, 1990.

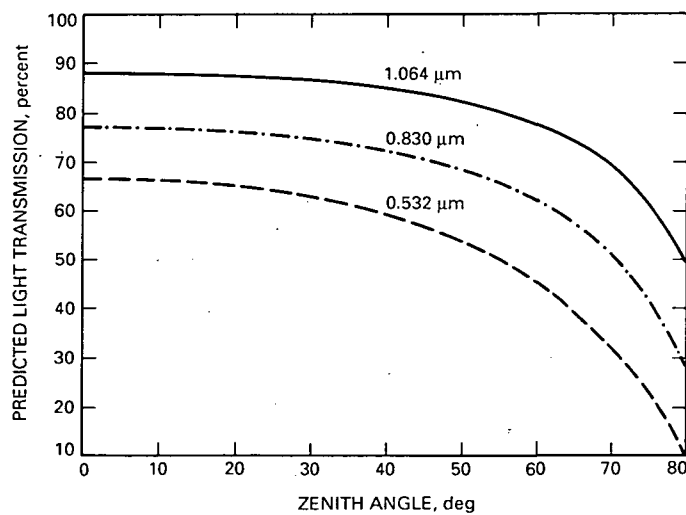


Fig. 10. Calculated light transmission at Goldstone for an average summer.

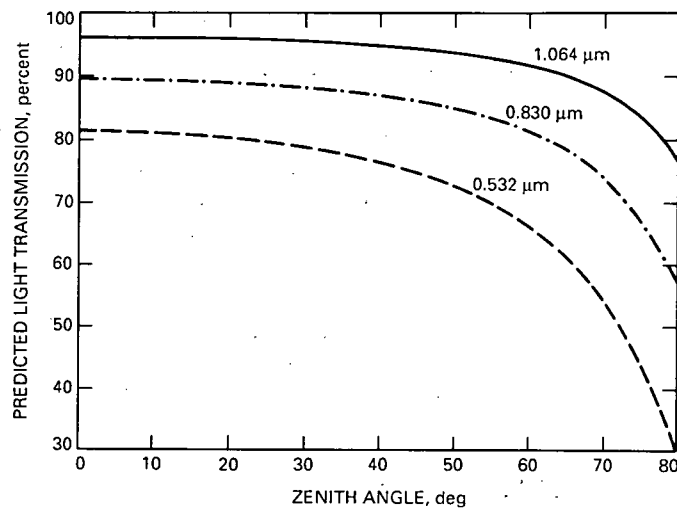


Fig. 11. Calculated light transmission at Table Mountain for an average summer.

442597

Modifications of the Griesmer Bound

R. J. McEliece¹ and G. Solomon²
Communications Systems Research Section

The Griesmer bound is a classical technique (developed in 1960) for estimating the minimum length n required for a binary linear code with a given dimension k and minimum distance d . In this article, a unified derivation of the Griesmer bound and two new variations on it are presented. The first variation deals with linear codes which contain the all-ones vector; such codes are quite common and are useful in practice because of their "transparent" properties. The second variation deals with codes that are constrained to contain a word of weight $\geq M$. In both cases these constraints (the all-ones word or a word of high weight) can increase the minimum length of a code with given k and d .

I. Introduction and Review of the Classical Griesmer Bound

The key notion for the Griesmer bound is what Solomon and Stiffler [8] called *puncturing*. If $x = (x_1, \dots, x_n)$ is a binary vector of length n and if $I \subseteq \{1, \dots, n\}$, the I -puncturing of x is the vector obtained by deleting the components of x indexed by I . Thus, for example, the $\{1, 4\}$ puncturing of (10101) is (011) . Puncturing is thus just a special kind of linear transformation, i.e., a projection onto certain coordinate positions, but here the traditional terminology will be retained. All of the results in this article, old and new, are based on the following simple combinatorial lemma.

Lemma. Let $a = (a_1, \dots, a_n)$ be a fixed binary vector of length n . If $b = (b_1, \dots, b_n)$ is another binary vector of

length n , let b' be the vector obtained by puncturing b at the positions where $a_i = 1$. Then

$$\text{wt}(b') = \frac{\text{wt}(b)}{2} + \frac{\text{wt}(a+b) - \text{wt}(a)}{2} \quad (1)$$

Proof: Without loss of generality, take

$$\left. \begin{array}{l} a = \overbrace{0000000}^{n-w} \overbrace{1111111}^w \\ b = 0001111 \ 1110000 \\ a+b = 0001111 \ 0001111 \end{array} \right\} \quad (2)$$

where $w = \text{wt}(a)$. Then, if $x = (x_1, x_2, \dots, x_n)$ is any vector of length n , then $x' = (x_1, \dots, x_{n-w})$. Similarly, define the complementary puncturing of x —at the components where $a_i = 0$ by $x'' = (x_{n-w+1}, \dots, x_n)$, so that

¹ Consultant, California Institute of Technology, Engineering Department.

² Consultant.

$\text{wt}(x) = \text{wt}(x') + \text{wt}(x'')$ for any vector x . Applying this rule to the second and third line of Eq. (2) yields, noting that $\text{wt}[(a+b)'] = \text{wt}(b')$ and $\text{wt}[(a+b)''] = w - \text{wt}(b'')$,

$$\text{wt}(b') + \text{wt}(b'') = \text{wt}(b)$$

$$\text{wt}(b') + [w - \text{wt}(b'')] = \text{wt}(a+b)$$

Adding these two equations, $2\text{wt}(b') = \text{wt}(b) + \text{wt}(a+b) - \text{wt}(a)$, which is the same as Eq. (1). \square

In the rest of the article, the MacWilliams-Sloane ([6], Section 1.1) terminology of an $[n, k, d]$ code is used to describe a binary linear code with length n , dimension k , and minimum distance d .

Theorem 1. Let C be an $[n, k, d]$ code, and let a be a codeword of weight d . Let C' be the code obtained from C by puncturing each codeword at the coordinates where $a_i = 1$. Then C' is an $[n-d, k-1, d']$ code with $d' \geq \lceil d/2 \rceil$.

Proof: The code C' is by definition of length $n-d$, since there are d punctured coordinates. To compute the dimension of C' , use the fact that the puncturing mapping P from C to C' is a linear transformation, so that $\text{rank}(P) + \text{nullity}(P) = \dim(C) = k$ ([4], Theorem 3.1.3). To find $\text{nullity}(P)$, examine the set of codewords $x \in C$ such that $x' = 0$. If x' is such a codeword, then the 1's of x must be confined to the d coordinates where a is nonzero, so that either $x = 0$ or $\text{wt}(a+x) < d$. But since $x+a$ is a codeword and d is the minimum weight of C , it follows that $x+a = 0$, i.e., $x = a$. Thus, there are just two words in C that, when punctured, yield 0—0 and a , and so $\text{nullity}(P) = 1$, so that $\text{rank}(P)$, i.e., the dimension of C' , is one less than the dimension of C , i.e., $k-1$. Finally, if b is an arbitrary codeword of C not equal to 0 or a , $\text{wt}(b+a) \geq d = \text{wt}(a)$, and so by the Lemma, $\text{wt}(b') \geq \lceil \text{wt}(b)/2 \rceil \geq \lceil d/2 \rceil$. Thus, every nonzero word in C' has weight $\geq \lceil d/2 \rceil$. \square

Let $n(k, d)$ be the minimum length of a binary code with Hamming distance $\geq d$ and dimension k . The original Griesmer bound can now be stated and proven ([3] or [6], p. 546).

Theorem 2 (Griesmer, 1960). If $k \geq 2$, then

$$n(k, d) \geq d + n(k-1, \lceil d/2 \rceil)$$

Proof: Let C be an $[n, k, d]$ binary linear code with $n = n(k, d)$. Then the code C' described in Theorem 1

is an $[n-d, k-1, d']$ code with $d' \geq \lceil d/2 \rceil$, and so its length must be $\geq n(k-1, \lceil d/2 \rceil)$. Hence, $n(k, d) - d \geq n(k-1, \lceil d/2 \rceil)$. \square

Corollary 1 (Griesmer, 1960).

$$n(k, d) \geq d + \lceil d/2 \rceil + \lceil d/2^2 \rceil + \dots + \lceil d/2^{k-1} \rceil \quad \text{for } k \geq 1$$

Proof: This follows from Theorem 2, combined with the self-evident result that $n(1, d) = d$ for all $d \geq 1$, with mathematical induction, and that $\lceil \lceil x \rceil / 2 \rceil = \lceil x/2 \rceil$ (see [1] or [5], solution to exercise 1.2.4, p. 476). \square

II. The Griesmer Bound for Codes Containing the All-Ones Word

In many applications, it is necessary to consider codes that contain the all-ones vector, e.g., "transparent codes" for synchronizing phase-shift-keyed-modulated data ([2], Section 6.6.1), or for synthesizing good finite state-codes [7]. It is therefore useful and interesting to study the possible loss in performance induced by requiring a code to contain the all-ones vector. Thus let $N(k, d)$ denote the minimum length of a binary code with Hamming distance $\geq d$ and dimension k that contains the all-ones vector.

Theorem 3. If $k \geq 2$, then (cf. Theorem 2).

$$N(k, d) \geq d + N(k-1, \lceil d/2 \rceil)$$

Proof: Let C be an $[n, k, d]$ binary linear code containing the all-ones vector with $n = N(k, d)$. Then the punctured code C' described in Theorem 2 is an $[n-d, k-1, d']$ code that contains the all-ones vector (since puncturing an all-ones vector leaves another all-ones vector) with $d' \geq \lceil d/2 \rceil$, and so its length must be $\geq N(k-1, \lceil d/2 \rceil)$. Hence, $N(k, d) - d \geq N(k-1, \lceil d/2 \rceil)$. \square

Theorem 4. Both $N(1, d) = d$ and $N(2, d) = 2d$ for all $d \geq 1$.

Proof: For the $k = 1$ result, take as the generator matrix

$$G = (1 \ 1 \ \dots \ 1)$$

For the $k = 2$ result, note that an $[n, 2, d]$ code with the all-ones vector has a $2 \times n$ generator matrix of the form

$$G = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 0 & 0 & \cdots & 1 & 1 \end{pmatrix}$$

Denote by n_0 the number of columns of G of the form $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and by n_1 the number of columns of the form $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Then, since the code has minimum weight d , it must follow that $n_1 \geq d$ and $n_0 \geq d$. Hence, $n = n_0 + n_1 \geq 2d$. On the other hand, by taking $n_0 = d$ and $n_1 = d$, one obtains a $[2d, 2, d]$ code containing the all-ones vector. \square

Theorem 5. If $k \geq 3$, then

$$N(k, d) \geq d + \lceil d/2 \rceil + \lceil d/2^2 \rceil + \cdots + 2\lceil d/2^{k-2} \rceil \quad (3)$$

Proof: This follows by mathematical induction on k , using Theorem 3 as the boundary value and Theorem 4 as the induction step, again with the help of the result $\lceil \lceil x \rceil / 2 \rceil = \lceil x/2 \rceil$ cited above. \square

Examples. Let $k = 3$ and $d = 3$. Then by the Corollary 1 and Theorem 5, $n(3, 3) \geq 3 + 2 + 1 = 6$ and $N(3, 3) \geq 3 + 2 + 2 = 7$. In both cases the bound is sharp, since there is a $[6, 3, 3]$ code, namely, a punctured $[7, 3, 4]$ simplex code with generator matrix

$$G = \begin{pmatrix} 110100 \\ 011010 \\ 001101 \end{pmatrix}$$

and a $[7, 3, 3]$ code with the all-ones word, namely,

$$G = \begin{pmatrix} 111111 \\ 100011 \\ 010010 \end{pmatrix}$$

Since $N(5, 9) \geq 9 + 5 + 3 + 2 \cdot 2 = 21$, there is no $[20, 5, 9]$ code with the all-ones word. There is, however, a $[21, 5, 9]$ code with the all-ones word, obtained from the $[16, 5, 8]$ biorthogonal code by repeating the information bits.

Theorem 6.

$$N(k, 2) = \begin{cases} k + 1 & \text{if } k \text{ is odd} \\ k + 2 & \text{if } k \text{ is even} \end{cases}$$

Proof: Since there is plainly no $[k, k, 2]$ code, with or without the all-ones word, it follows that $N(k, 2) \geq k + 1$ for all k . The only $[k + 1, k, 2]$ code has the parity-check matrix

$$H = \begin{pmatrix} \overbrace{11 \cdots 1}^{k+1} \\ 11 \cdots 1 \end{pmatrix}$$

This code contains the all-ones vector if and only if k is odd, which proves that $N(k, 2) = k + 1$ if k is odd, and $N(k, 2) \geq k + 2$ if k is even. If k is even, there is a $[k + 2, k, 2]$ code containing the all-ones word, with a parity-check matrix (illustrated for $k = 6$)

$$H = \begin{pmatrix} 1111111 \\ 1100000 \end{pmatrix}$$

so that $N(k, 2) = k + 2$ when k is even, as asserted. \square

Corollary 2.

$$N(k, 3) \geq \begin{cases} k + 3 & \text{if } k \text{ is even} \\ k + 4 & \text{if } k \text{ is odd} \end{cases}$$

Proof: From Theorem 3, $N(k, 3) \geq 3 + N(k - 1, 2)$. The result now follows from Theorem 6. \square

III. The Griesmer Bound for Codes Containing a Word of Bounded Weight

As another variation on the Griesmer bound, let $N(k, d, M)$ denote the length of the shortest $[n, k, d]$ binary linear code that contains a word of weight $\geq M$.

Theorem 7.

$$N(k, d, M) \geq d + N(k - 1, \lceil d/2 \rceil, \lceil M/2 \rceil)$$

Proof: Let C be an $[n, k, d]$ code containing a word of weight $\geq M$. As in the proof of Theorem 2, consider the code C' , which is an $[n - d, k - 1, d']$ code with $d' \geq \lceil d/2 \rceil$. Now let b be a word of weight $\geq M$ in C . Then, by the Lemma, $\text{wt}(b') \geq \lceil \text{wt}(b)/2 \rceil \geq \lceil M/2 \rceil$. Thus, C' is an $[n - d, k - 1, d']$ code with $d' \geq \lceil d/2 \rceil$ containing a word of weight $\geq \lceil M/2 \rceil$, i.e., $n - d \geq N(k - 1, \lceil d/2 \rceil, \lceil M/2 \rceil)$. \square

Theorem 8. If $M \geq d$ and $k \geq 2$,

$$N(k, d, M) \geq d + \lceil d/2 \rceil + \lceil d/2^2 \rceil + \dots$$

$$+ \lceil d/2^{k-2} \rceil + \lceil M/2^{k-1} \rceil$$

Proof: This follows from Theorem 3 and the boundary value $N(1, d, M) = \max(M, d)$. \square

Example. According to Theorem 5, $n(3, 4) \geq 7$, and there is a $[7, 3, 4]$ code, i.e., the simplex code. However, this code has words only of weight 4. If one looks for a

$[7, 3, 4]$ code with a word of weight 5 or more, an appeal to Theorem 8 shows that $N(3, 4, 5) \geq 4 + \lceil 4/2 \rceil + \lceil 5/4 \rceil = 8$. There is an $[8, 3, 4]$ code with a word of weight 6, namely, the code with generator matrix

$$G = \begin{pmatrix} 11111100 \\ 00001111 \\ 11001010 \end{pmatrix}$$

but it is unknown whether there is an $[8, 3, 4]$ code with a word of weight 5.

References

- [1] L. D. Baumert and R. J. McEliece, "A Note on the Griesmer Bound," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 134-135, January 1973.
- [2] G. C. Clark, Jr. and J. B. Cain, *Error-Correction Coding for Digital Communications*, New York: Plenum Press, 1981.
- [3] J. H. Griesmer, "A Bound for Error-Correcting Codes," *IBM J. Res. Develop.*, vol. 4, pp. 532-542, 1960.
- [4] K. Hoffman and R. Kunze, *Linear Algebra*, Englewood Cliffs, New Jersey: Prentice-Hall, 1961.
- [5] D. E. Knuth, *The Art of Computer Algorithms*, vol. 1, Reading, Massachusetts: Addison-Wesley, p. 476, 1973.
- [6] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, Amsterdam: North-Holland, 1977.
- [7] F. Pollara, R. J. McEliece, and K. Abdel-Ghaffar, "Finite-State Codes," *IEEE Trans. Inform. Theory*, vol. IT-34, pp. 1083-1088, September 1988.
- [8] G. Solomon and J. J. Stiffler, "Algebraically Punctured Cyclic Codes," *Information and Control*, vol. 8, pp. 170-179, 1965.

Integer Cosine Transform for Image Compression

K.-M. Cheung, F. Pollara, and M. Shahshahani
Communications Systems Research Section

This article describes a recently introduced transform algorithm called the integer cosine transform (ICT), which is used in transform-based data compression schemes. The ICT algorithm requires only integer operations on small integers and at the same time gives a rate-distortion performance comparable to that offered by the floating-point discrete cosine transform (DCT). The article addresses the issue of implementation complexity, which is of prime concern for source coding applications of interest in deep-space communications. Complexity reduction in the transform stage of the compression scheme is particularly relevant, since this stage accounts for most (typically over 80 percent) of the computational load.

I. Introduction

The rate-distortion performance of three transform-based coding schemes used to compress the test images for the Comet Rendezvous Asteroid Flyby (CRAF)/Cassini Project was presented in [1]. More recently, the issue of implementation complexity, which is of prime concern to spacecraft applications, was addressed. The computational bottleneck of transform-based algorithms lies in the front-end transform stage, which accounts for over 80 percent of the computational load of these compression schemes. This article describes a recently introduced transform algorithm called the integer cosine transform (ICT), which requires only integer operations on small integers and at the same time has rate-distortion comparable to that of the floating-point discrete cosine transform (DCT), which is the most practical and near optimal approach known for data compression. The implementation complexity of the ICT is substantially lower than that of the DCT, and is comparable to that of the Hadamard transform (HT). The ICT is a practical approach to achieving the high-rate deep-space communications that are possible with the DCT.

II. Background: Transform-Based Schemes

In preparing the test images for the CRAF/Cassini Project, three transform-based encoding algorithms were used to compress a set of seven planetary images [1], which are continuous-tone gray-scale, with pixel values ranging from 0 (black) to 255 (white). All three algorithms can be viewed as consisting of three stages, as illustrated in Fig. 1: the data transform stage, the quantization stage, and the entropy-coding stage. The compression algorithms work on a block-by-block basis, i.e., they compress an 8×8 block of the picture at a time. In each algorithm, the encoder first applies an 8×8 floating-point DCT or an 8×8 HT to the picture block to generate an 8×8 block of transform coefficients. These numbers are then quantized by a predetermined 8×8 quantization template to integer values. Most quantized values have small magnitudes. Due to the skewed distribution of the quantized transform coefficients, compression is achieved by assigning shorter transmission-bit patterns to the more frequently occurring integers. This is realized in the last stage of the compression scheme, the entropy coder, which maps the quantized values to appropriate transmission-bit patterns.

In the CRAF/Cassini data compression experiment, all three transform-based schemes used the same DCT or HT (stage 1) and the same quantization template (stage 2). The difference lies in the choice of entropy coder in the third stage, where one may use the Joint Photographic Expert Group (JPEG) Huffman code [2,12], an arithmetic code [2], or the Gallager-van Voorhis-Huffman (GVH) code [3]. In general, the DCT-based schemes are more effective (0.1 to 0.3 bits per pixel) than the HT-based schemes, especially in the high bit rate (near lossless) range. However, the more effective DCT-based schemes are more computationally intensive than the HT-based schemes. The major computational burden of DCT-based schemes lies in the DCT stage, which requires a large number of floating-point multiplications and additions. HT-based schemes, on the other hand, require only integer additions and subtractions in the transform stage. From the hardware's point of view, floating-point operations are much slower and more difficult to implement than the corresponding integer operations. For a general N -point DCT, a straightforward algorithm [4] that yields a simple regular implementation and a small chip size requires $2N^3$ multiplications and $2N^3$ additions. A more sophisticated N -point fast DCT [5], where N has a power of 2, that uses complex data-shuffling strategies still requires $N^2 \log_2 N$ multiplications and $N(3N \log_2 N - N + 1)$ additions. The large number of floating-point operations required to perform DCT, particularly for large N , is the computational bottleneck for all DCT-based signal-processing schemes.

III. Integer Cosine Transform

Recently Choy, Cham, and Lee [6] proposed a new 8-point transform called the integer cosine transform (ICT), which requires only integer multiplications and additions, and thus is much simpler to implement than the DCT. An ICT chip was fabricated and was proven to be efficient in both silicon area and speed [6]. The 8×8 ICT matrix suggested in [6] is given in Fig. 2(c). Notice that the elements in the matrix are all integers, and the ICT matrix B in Fig. 2(b) has sign and magnitude patterns that resemble those of the DCT matrix A in Fig. 2(a). The similarity of the ICT matrix to the DCT matrix, together with the orthogonality property of the ICT ($BB^T = \Delta$, where Δ is a diagonal matrix), guarantees that the ICT, as well as its inverse, possesses the same transform structure as the DCT. Thus, any fast DCT algorithms can be used to compute a fast ICT.

This 8×8 ICT matrix was used to compute a two-dimensional 8×8 transform and then compress the planetary images *saturn1* and *saturn2*. The transform coefficients

were quantized by using the same quantization template as in the aforementioned DCT-based schemes. The entropy of the quantized transform coefficients and the mean square error (MSE) of the reconstructed picture were computed, and the results are shown in Fig. 3. These simulation results indicate that any difference in rate-distortion performance resulting from using the floating-point DCT or the ICT is unnoticeable.

Although the 8-point ICT proposed by Choy, Cham, and Lee performs remarkably well, it is quite ad hoc, and no general mathematical formulation of the ICT is given in [6]. The contributions of this article are to put the ICT into a more formal mathematical setting, and to generalize their idea to any N -point ICT. The mathematical properties of the ICT are investigated in the following sections. Since the ICT is separable, and the extension of the one-dimensional ICT to two dimensions is straightforward, this article focuses on the one-dimensional case. Section IV gives a characterization of ICT matrices. An 8×8 ICT matrix that is multiplication-free and requires only binary additions and shifts is given in Section V (the MSE versus entropy performance of the multiplication-free ICT, the original ICT of [6], and that of the floating-point DCT are shown in Fig. 3). A general procedure for the construction of an $N \times N$ ICT matrix is obtained in Section VI; and two 16×16 ICT matrices, one with only small integer entries and one with all entries' powers of two (multiplication-free) are exhibited in Section VII.

IV. Mathematical Properties of the ICT

The integer cosine transform and the discrete cosine transform are closely related. Let C and A be the respective ICT and DCT $N \times N$ matrices. An orthonormal matrix (i.e., $AA^T = I$), $A = [a_{kn}]$, is defined as follows for $0 \leq n \leq N-1$:

$$\begin{aligned} a_{kn} &= \frac{1}{\sqrt{N}} & k &= 0 \\ &= \frac{2}{N} \cos \frac{\pi(2n+1)k}{2N} & 1 \leq k \leq N-1 \end{aligned} \quad (1)$$

Using A as a template, the ICT matrix $C = [c_{kn}]$ is an orthogonal matrix (i.e., $CC^T = \Delta$, where Δ is a diagonal matrix) with the following properties:

- (1) Integer property: c_{kn} represents integers for $0 \leq k, n \leq N-1$.
- (2) Orthogonality property: Rows (or columns) of C are orthogonal.

(3) Relationship with DCT:

- (a) $\text{sgn}(c_{kn}) = \text{sgn}(a_{kn})$ for $0 \leq k, n \leq N-1$.
- (b) If $a_{kn} = a_{st}$, then $c_{kn} = c_{st}$ for $0 \leq k, n, s, t \leq N-1$.

The integer property eliminates real multiplication and real addition operations. The orthogonality property assures that the inverse ICT has the same transform structure as the ICT. Notice that C is only required to be orthogonal, but not orthonormal. However, any orthogonal matrix can be made orthonormal by multiplying it by an appropriate diagonal matrix. This operation can be incorporated in the quantization (dequantization) stage of the compression (decompression) scheme, thus sparing the ICT (inverse ICT) from floating-point operations, and at the same time preserving the same transform structure as in the floating-point DCT (inverse DCT). The relationship between ICT and DCT guarantees efficient energy packing and allows the use of any fast DCT technique for the ICT.

V. ICT for $N = 8$

The floating-point 8×8 DCT matrix is shown in Fig. 2(a). A general structure of the 8×8 ICT matrix is given in Fig. 2(b). The symbols a, b, c, d, e , and f in Fig. 2(b) are numbers that satisfy conditions (1) through (3) given in Section IV. It was suggested in [6] to use $a = 5, b = 3, c = 2, d = 1, e = 3$, and $f = 1$ for the $N = 8$ ICT, as shown in Fig. 2(c). There are many other sets of (a, b, c, d, e, f) that can generate an orthogonal ICT. The integer set (a, b, c, d, e, f) gives an orthogonal ICT if, and only if, CC^t is a diagonal matrix. This is equivalent to the requirement

$$ab - ac - bd - cd = 0 \quad (2)$$

with e and f arbitrary. The integer set $(4, 2, 2, 0, 4, 2)$ satisfies Eq. (2) and the corresponding ICT matrix is given in Fig. 2(d). Notice that the integers chosen are all powers of 2, and thus only simple binary additions and shift operations are required for this ICT. The MSE versus entropy performance of the compression scheme using this multiplication-free ICT is shown in Fig. 3. In view of the particular choice of the integers in the multiplication-free implementation, one expects the performance of this ICT to be inferior to that of the floating-point DCT and the ICT of [6]. However, the difference in the performance is small, as shown in Fig. 3.

VI. A General Procedure for Constructing an ICT Matrix

A general procedure to construct an $N \times N$ ICT matrix is presented in this section. For any $N \times N$ ICT matrix, this construction is done on the ground prior to implementation. The DCT matrix is used as a template to generate an ICT matrix. The procedure is as follows:

- (1) Generate the $N \times N$ DCT matrix A .
- (2) Construct an $N \times N$ matrix B by substituting the N possible absolute values in A with N symbols, and preserve the signs of the elements in A .
- (3) Evaluate BB^t , and generate a set of independent algebraic equations that force BB^t to be a diagonal matrix.
- (4) Find a set of N numbers that satisfy the set of algebraic equations generated in (3).

Since for a given N , there are $N(N-1)$ nondiagonal elements in C , part (3) of the procedure gives $N(N-1)/2$ quadratic equations. This set of equations is too large to be handled easily except for small N . However, by setting the most frequently occurring symbol in C to be an integer such as 1 or 2, the number of independent equations decreases substantially. As shown above, when $N = 8$, the number of equations is reduced from 28 to 1. The most tedious part of the above procedure is part (4), that is, finding N integers that satisfy the set of nonlinear algebraic equations generated in part (3). By using such advanced symbolic manipulation tools as *Mathematica* [7], the effort of generating the set of algebraic equations in part (3) and solving them in part (4) can be greatly reduced. In fact, *Mathematica* was used in an interactive manner to generate the 8×8 and 16×16 ICT matrices given in this article.

In order to obtain good compression performance, the set of $N-1$ integers must have a magnitude profile similar to the $N-1$ floating-point elements of A . Furthermore, if the multiplication-free property is desired, the set of N integers must be restricted to powers of 2. Some ad hoc techniques are usually needed to simplify the above calculations.

Note also that there is a general procedure for approximating an orthonormal matrix arbitrarily closely to one with rational coefficients. Given an orthonormal matrix Q with no eigenvalue equal to -1 , let $S = (I - Q)(I + Q)^{-1}$. Then S is skew symmetric since

$$S + S^t = (I - Q)(I + Q)^{-1} + (I + Q^t)^{-1}(I - Q^t)$$

$$= (I - Q)(I + Q)^{-1} + (I + Q^{-1})^{-1}(I - Q^{-1})$$

$$= (I - Q)(I + Q)^{-1} + (I + Q)^{-1}(Q - I)$$

$$= 0 \quad (3)$$

Conversely, if S is any skew-symmetric matrix such that -1 is not an eigenvalue, then by essentially the same computation, the matrix $Q = (I - S)(I + S)^{-1}$ is orthonormal. Thus, given an orthonormal matrix Q , one can approximate $S = (I - Q)(I + Q)^{-1}$ by an arbitrarily close rational skew-symmetric matrix S' . Then, $Q' = (I - S')(I + S')^{-1}$ is a rational orthonormal matrix close to Q . While this procedure works well in theory, there are practical difficulties in its application. In practice, one considers $\lambda Q'$ where λ is an integer so that $\lambda Q'$ is integral. The matrix $\lambda Q'$ obtained by this procedure has generally large entries, which makes it unsuitable for many applications.

VII. ICT for $N = 16$

In this section, the general procedure of Section VI is used to construct a 16×16 ICT matrix. From Eq. (1), one obtains the 16×16 DCT matrix A . Notice that there are 16 non-negative values in A . The 16×16 ICT matrix B shown in Fig. 4(a) is obtained by using A as a template. By setting $a = 1$ and forcing all nondiagonal elements in BB^t to be zero, one obtains the following set of four independent nonlinear equations:

$$bc + cf - df - bg - eg - eh + di - hi = 0 \quad (4)$$

$$bd - ce - de + fg + bh - fh + ci + gi = 0 \quad (5)$$

$$-cd + be + bf - cg - dh + gh + ei - fi = 0 \quad (6)$$

$$jk - jl - km - lm = 0 \quad (7)$$

Notice that n and o are arbitrary. By extensive search on the sets of numbers that satisfy Eqs. (4), (5), (6), and (7), the following solution, which has a magnitude profile similar to that of the $N - 1$ floating-point elements of the 16×16 DCT matrix, was obtained: $a = 1, b = 18, c = 18, d = 16, e = 14, f = 14, g = 7, h = 10, i = 2, j = 10, k = 9, l = 6, m = 2, n = 56$, and $o = 2$. The corresponding ICT matrix is shown in Fig. 4(b). Another solution to the above system of equations is $a = 1, b = 4, c = 4, d = 0, e = 2, f = 2, g = 4, h = 0, i = 0, j = 4, k = 2, l = 2, m = 0, n = 1$, and $o = 4$. This matrix is shown in Fig. 4(c). Notice that all integers in this solution are powers of 2, so only binary shift and addition operations are required for this ICT transform. Since there are many zeros in this solution, one does not expect it to give an efficient ICT matrix. Intuitively, a transform matrix with good energy compaction should not have many zeros.

VIII. Conclusion

This article explored the mathematical properties of a new class of integer transforms called the integer cosine transform and derived a general construction procedure for this transform. This procedure can be used to construct integer versions of other transforms, such as the Fourier [8], sine [9], Gabor [10], wavelet [11], and so forth. The basic idea is to approximate a floating-point transform with its integer counterpart in the hope of achieving comparable performance with much-reduced implementation complexity. In the case of the discrete cosine transform, its integer counterpart, the ICT, has an implementation complexity substantially lower than that of the DCT and comparable to that of the Hadamard transform. Simulation results indicate that rate-distortion performance of the ICT is only slightly inferior to that of the DCT.

Acknowledgment

The authors thank I. Onyszchuk of the Jet Propulsion Laboratory for his helpful discussion.

References

- [1] S. J. Dolinar, K.-M. Cheung, I. Onyszchuk, F. Pollara, and S. Arnold, "Compressed/Reconstructed Test Images for CRAF/Cassini," *TDA Progress Report 42-104*, vol. October–December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 88–97, February 15, 1991.
- [2] *Joint Photographic Experts Group (JPEG) Draft Technical Specification (Rev. 5)*, International Standards Organization/International Consultative Committee for Telephone and Telegraph (ISO/CCITT), Washington, D.C., January 15, 1990.
- [3] K.-M. Cheung, P. Smyth, and H. Wang, "A High-Speed Distortionless Predictive Image-Compression Scheme," *TDA Progress Report 42-102*, vol. April–June 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 73–90, August 15, 1990.
- [4] N. Ahmed, T. Natarajan, and K. Rao, "Discrete Cosine Transform," *IEEE Trans. Comput.*, vol. C-23, no. 1, pp. 90–93, January 1974.
- [5] M. Narasimha and A. Peterson, "On the Computation of the Discrete Cosine Transform," *IEEE Trans. Comm.*, vol. COM-26, no. 6, pp. 934–946, June 1978.
- [6] C. Choy, W. Cham, and L. Lee, "An Integer Cosine Transform Chip Using ASIC Technology," *Proceedings of the Picture Coding Symposium*, Torino, Italy, pp. 5.5.1–5.5.2, September 1988.
- [7] S. Wolfram, *Mathematica: A System for Doing Mathematics by Computer*, Redwood City, California: Addison-Wesley Publishing Company, 1988.
- [8] J. Cooley and J. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," *Math. of Comput.*, vol. 19, no. 90, pp. 297–301, April 1965.
- [9] P. Yip and K. Rao, "A Fast Computational Algorithm for the Discrete Sine Transform," *IEEE Trans. Comm.*, vol. COM-28, no. 2, pp. 304–307, February 1980.
- [10] D. Gabor, "Theory of Communications," *Institution of Electrical Engineering Journal*, vol. 93, no. 22, pp. 429–457, March 1946.
- [11] A. Grossmann and J. Morlet, "Decomposition of Hardy Functions Into Square Integrable Wavelets of Constant Shape," *SIAM J. Math. Anal.*, vol. 15, no. 4, pp. 723–736, July 1984.
- [12] F. Pollara and S. Arnold, "Emerging Standards for Still-Image Compression: A Software Implementation and Simulation Study," *TDA Progress Report 42-104*, vol. October–December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 98–102, February 15, 1991.

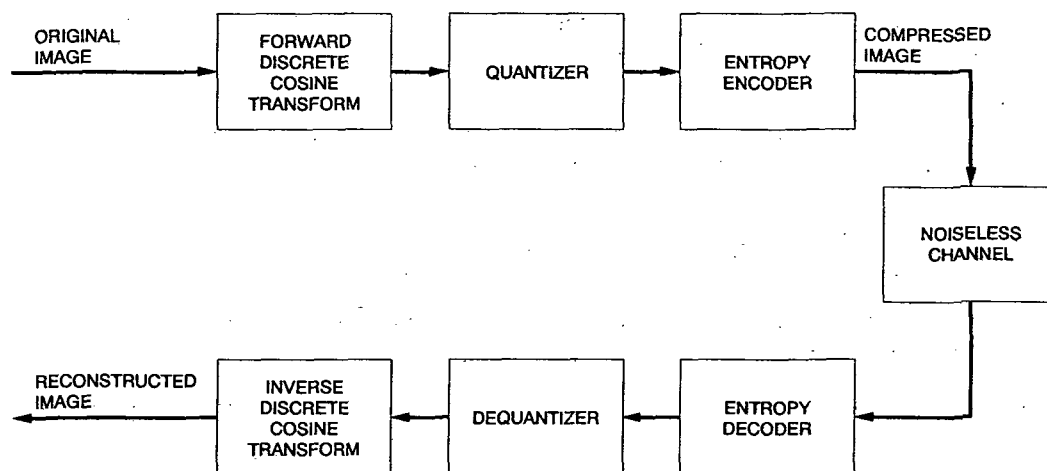


Fig. 1. DCT-based compression system.

(a)	$A = \begin{bmatrix} 0.354 & 0.354 & 0.354 & 0.354 & 0.354 & 0.354 & 0.354 & 0.354 \\ 0.490 & 0.416 & 0.278 & 0.098 & -0.098 & -0.278 & -0.416 & -0.490 \\ 0.462 & 0.191 & -0.191 & -0.462 & -0.462 & -0.191 & 0.191 & 0.462 \\ 0.416 & -0.098 & -0.490 & -0.278 & 0.278 & 0.490 & 0.098 & -0.416 \\ 0.354 & -0.354 & -0.354 & 0.354 & 0.354 & -0.354 & -0.354 & 0.354 \\ 0.278 & -0.490 & 0.098 & 0.416 & -0.416 & -0.098 & 0.490 & -0.278 \\ 0.191 & -0.462 & 0.462 & -0.191 & -0.191 & 0.462 & -0.462 & 0.191 \\ 0.098 & -0.278 & 0.416 & -0.490 & 0.490 & -0.416 & 0.278 & -0.098 \end{bmatrix}$
(b)	$B = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ a & b & c & d & -d & -c & -b & -a \\ e & f & -f & -e & -e & -f & f & e \\ b & -d & -a & -c & c & a & d & -b \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ c & -a & d & b & -b & -d & a & -c \\ f & -e & e & -f & -f & e & -e & f \\ d & -c & b & -a & a & -b & c & -d \end{bmatrix}$
(c)	$C = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 3 & 2 & 1 & -1 & -2 & -3 & -5 \\ 3 & 1 & -1 & -3 & -3 & -1 & 1 & 3 \\ 3 & -1 & -5 & -2 & 2 & 5 & 1 & -3 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 2 & -5 & 1 & 3 & -3 & -1 & 5 & -2 \\ 1 & -3 & 3 & -1 & -1 & 3 & -3 & 1 \\ 1 & -2 & 3 & -5 & 5 & -3 & 2 & -1 \end{bmatrix}$
(d)	$D = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 4 & 2 & 2 & 0 & 0 & -2 & -2 & -4 \\ 4 & 2 & -2 & -4 & -4 & -2 & 2 & 4 \\ 2 & 0 & -4 & -2 & 2 & 4 & 0 & -2 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 2 & -4 & 0 & 2 & -2 & 0 & 4 & -2 \\ 2 & -4 & 4 & -2 & -2 & 4 & -4 & 2 \\ 0 & -2 & 2 & -4 & 4 & -2 & 2 & 0 \end{bmatrix}$

Fig. 2. Four 8×8 transform matrices: (a) DCT matrix (real number entitles rounded off to nearest thousandth); (b) structure of ICT matrix; (c) ICT matrix from [6]; and (d) multiplication-free ICT.

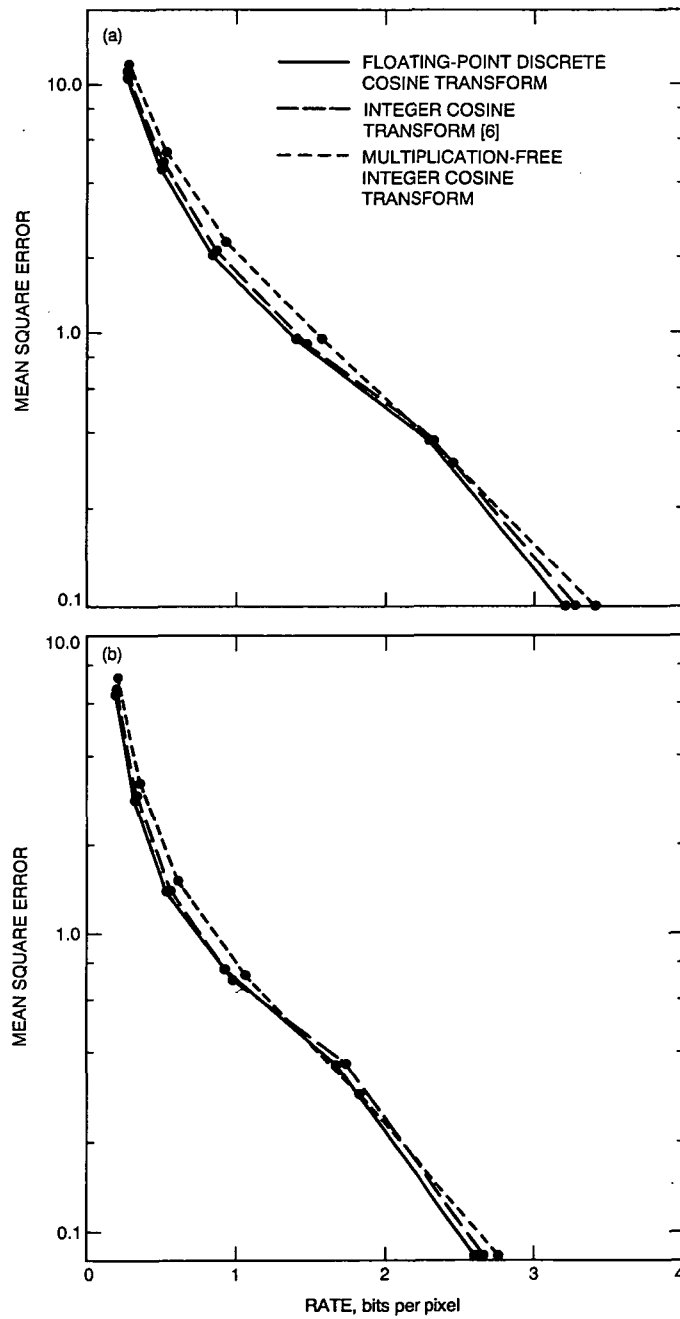


Fig. 3. Rate-distortion performance of three transform algorithms: (a) image *saturn1*, and (b) image *saturn2*.

(a)

a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a
b	c	d	e	f	g	h	i	-i	-h	-g	-f	-e	-d	-c	-b
j	k	l	m	-m	-l	-k	-j	-j	-k	-l	-m	m	l	k	j
c	f	i	-g	-d	-b	-e	-h	h	e	b	d	g	-i	-f	-c
o	n	-n	-o	-o	-n	n	o	o	n	-n	-o	-o	-n	n	o
d	i	-e	-c	-h	f	b	g	-g	-b	-f	h	c	e	-i	-d
k	-m	-j	-l	l	j	m	-k	-k	m	j	l	-l	-j	-m	k
e	-g	-c	i	b	h	-d	-f	f	d	-h	-b	-i	c	g	-e
a	-a	-a	a	a	-a	-a	a	a	-a	-a	a	a	-a	-a	a
f	-d	-h	b	-i	-c	g	e	-e	-g	c	i	-b	h	d	-f
l	-j	m	k	-k	-m	j	-l	-l	j	-m	-k	k	m	-j	l
g	-b	f	h	-c	e	i	-d	d	-i	-e	c	-h	-f	b	-g
n	-o	o	-n	-n	o	-o	n	n	-o	o	-n	-n	o	-o	n
h	-e	b	-d	g	i	-f	c	-c	f	-i	-g	d	-b	e	-h
m	-l	k	-j	j	-k	l	-m	-m	l	-k	j	-j	k	-l	m
i	-h	g	-f	e	-d	c	-b	b	-c	d	-e	f	-g	h	-i

(b)

1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	18	16	14	14	7	10	2	-2	-10	-7	-14	-14	-16	-18	-18
10	9	6	2	-2	-6	-9	-10	-10	-9	-6	-2	2	6	9	10
18	14	2	-7	-16	-18	-14	-10	10	14	18	16	7	-2	-14	-18
2	5	-5	-2	-2	-5	5	2	2	5	-5	-2	-2	-5	5	2
16	2	-14	-18	-10	14	18	7	-7	-18	-14	10	18	14	-2	-16
9	-2	-10	-6	6	10	2	-9	-9	2	10	6	-6	-10	-2	9
14	-7	-18	2	18	10	-16	-14	14	16	-10	-18	-2	18	7	-14
1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1
14	-16	-10	18	-2	-18	7	14	-14	-7	18	2	-18	10	16	-14
6	-10	2	9	-9	-2	10	-6	-6	10	-2	-9	9	2	-10	6
7	-18	14	10	-18	14	2	-16	16	-2	-14	18	-10	-14	18	-7
5	-2	2	-5	-5	2	-2	5	5	-2	2	-5	-5	2	-2	5
10	-14	18	-16	7	2	-14	18	-18	14	-2	-7	16	-18	14	-10
2	-6	9	-10	10	-9	6	-2	-2	6	-9	10	-10	9	-6	2
2	-10	7	-14	14	-16	18	-18	18	-18	16	-14	14	-7	10	-2

(c)

2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
4	4	0	2	2	4	0	0	0	0	-4	-2	-2	0	-4	-4
4	2	2	0	0	-2	-2	-4	-4	-2	-2	0	0	2	2	4
4	2	0	-4	0	-4	-2	0	0	2	4	0	4	0	-2	-4
4	1	-1	-4	-4	-1	1	4	4	1	-1	-4	-4	-1	1	4
0	0	-2	-4	0	2	4	4	-4	-4	-2	0	4	2	0	0
2	0	-4	-2	2	4	0	-2	-2	0	4	2	-2	-4	0	2
2	-4	-4	0	4	0	0	-2	2	0	0	-4	0	4	4	-2
2	-2	-2	2	2	-2	-2	2	2	-2	-2	2	2	-2	-2	2
2	0	0	4	0	-4	4	2	-2	-4	4	0	-4	0	0	-2
2	-4	0	2	-2	0	4	-2	-2	4	0	-2	2	0	-4	2
4	-4	2	0	-4	2	0	0	0	0	-2	4	0	-2	4	-4
1	-4	4	-1	-1	4	-4	1	1	-4	4	-1	-1	4	-4	1
0	-2	4	0	4	0	-2	4	-4	2	0	-4	0	-4	2	0
0	-2	2	-4	4	-2	2	0	0	2	-2	4	-4	2	-2	0
0	0	4	-2	2	0	4	-4	4	-4	0	-2	2	-4	0	0

Fig. 4. Three 16×16 ICT matrices: (a) structure of ICT matrix; (b) ICT matrix; and (c) multiplication-free ICT matrix.

1994025117

N94-29620

TDA Progress Report 42-105

442603

May 15, 1991

A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation

Part II: Tests on the Antenna

T. Y. Otoshi, S. R. Stewart, and M. M. Franco
Ground Antennas and Facilities Engineering Section

Part I described an X-band (8.45 GHz) test package for testing the new 34-m beam-waveguide antenna at Goldstone. In addition, results were given for the test package in an "on-the-ground" configuration. This article gives results for the test package in an "on-the-antenna" configuration. Included are X-band zenith noise-temperature values and tipping-curve data obtained at the Cassegrain focal point F1 as well as at the pedestal room focal point F3. Subreflector Z-defocus test results for both F1 and F3 are also presented. The X-band test package operated well in all of the different test configurations and exceeded expected performance.

I. Introduction

As described in Part I [1], a test package has been developed to test the new 34-m beam-waveguide (BWG) antenna at X-band (8.45 GHz). The portable test package can be transported to different focal point locations of the BWG system. The degradations caused by the BWG mirror systems are determined by differencing the measured parameters at the different focal points (Fig. 1).

Part I gave the results of noise-temperature measurements for the X-band test package on the ground at 8.45 GHz. This article presents the results of zenith operating system noise-temperature measurements as well as results from tipping-curve and subreflector tests on the new BWG antenna at the Cassegrain focal point F1 and at the pedestal room focal point F3.

Noise-temperature symbols are used in many of the tables in this article. For the reader's convenience, the symbols are defined in Table 1.

II. Installations

Figure 2 shows the X-band 29-dBi horn test package installed on the antenna at F1. A removable wooden floor (not shown) about 12 inches below the test package and a ladder (shown in Fig. 2) are parts of a temporary installation. This installation facilitates ease of connecting cables and servicing the test package. The structure shown supporting the X-band test package is a universal mount that allows any of the three 29-dBi horn configuration test packages (X-band, 8.45 GHz; Ku-band, 12 GHz; and Ka-band, 32 GHz) to be interchanged and installed. A tape

measure showed that the desired and actual horn phase center locations agreed to within 1/4 to 1/8 in., with measurements accurate to $\pm 1/16$ in. Any error of up to 2 or 3 inches in the phase center location in the Z direction can be compensated for by readjustment of the subreflector.

After completion of noise-temperature and antenna-efficiency measurements at F1, the X-band test package was removed, reconfigured to a 22-dBi horn configuration and installed at F3 (Fig. 3). The mounting table shown in Fig. 3 is a universal mount that can support any of the test packages (X-, Ku-, or Ka-band) and provides three-axis adjustment of the test package location. Adjustments of ± 3 in. can be made along three orthogonal axes (i.e., the vertical direction, the radial direction towards and away from the hub center, and the other transverse direction).

III. Test Results

A. Preliminary Diagnostics Work

When the X-band test package was first installed at F1 and the new BWG antenna was tested, the zenith noise temperature was about 28 K as compared to an expected value of about 26 K. Various tests were run to isolate contributors to the high noise-temperature values. Six contributors were found and are listed in Table 2. With the exception of noise contributed from the open hub area below the feed horn, all noise contributions from various sources were minimized or eliminated by covering openings with perforated sheets or aluminum tape.

B. Zenith Noise-Temperature Measurements

Figures 4(a) and (b) show typical mini-cal [1] plots with the 29-dBi horn test package mounted at F1 and with all of the contributors of Table 2 eliminated except item 6. Figure 5 shows a mini-cal plot of the operating system temperature when the test package was mounted at F3. For these plots, corrections were made for gain changes but no corrections were made for nonlinearity, which was typically less than 1 percent. Since the uncertainty of the linearity corrections is itself about ± 1 percent, the small linearity corrections, if made, would not significantly enhance the quality of the system noise-temperature data shown on these particular plots.

A summary of results of X-band zenith operating system noise temperatures for the ground and the F1 and F3 locations is given in Table 3. The value shown in the table for each test configuration is the grand average of the average operating system temperature determined for the individual observing periods. Observations were made

over a 7-month period. Each grand average value shown in Table 3 is estimated to have an uncertainty of ± 0.5 K (one standard deviation).

Although some F2 measurements were performed for diagnostic purposes, they are not included in Table 3. The F2 measurements were questionable due to uncertainties as to whether or not a low-noise diode signal from the noise box (Fig. 6) had been left on during the F2 testing.

When the test package is on the ground, the general expression for the operating system temperature is

$$T_{op} = T'_{cb}/(L_{atm}L_{wg}) + T_{atm}/L_{wg} + T_{wg} + T_{hemt} + T_{fup} \quad (1)$$

Under standard conditions at 8.45 GHz, the component values are $T'_{cb} = 2.5$ K, $T_{atm} = 2.17$ K, $T_{wg} = 4.69$ K, $T_{hemt} = 13$ K, $T_{fup} = 0.4$ K, $L_{atm} = 1.00814$ (corresponding to 0.0352 dB), and $L_{wg} = 1.0163$ (corresponding to 0.07 dB).

Substitutions of the values into Eq. (1) result in a predicted T_{op} of 22.7 K, which agrees with the measured ground value of 22.7 K shown in Table 3.

Differential noise temperatures for the various configurations are shown in Table 4. The F1-minus-ground-configuration data have a value of 3.2 K, which is due to subreflector scattering, tripod scattering, leakage through the main reflector surface, and main reflector spillover. This value can be compared to an F1-minus-ground-configuration value of 3.8 K for X-band reported for the DSS 15 34-m Cassegrain antenna.¹ The F3-minus-F1-configuration data noise-temperature values provide a measure of the degradation caused by the six mirrors. A slight increase of 0.6 K in the 8.45-GHz noise temperature occurred after the ellipsoid and other five mirrors were realigned for purposes of improving the F3 BWG antenna performance at 32 GHz.

Table 5 is a worksheet showing how the final values summarized in Table 3 were derived. Shown are the observation periods, the measured operating system noise-temperature values, weather information, and normalized values after corrections were made for weather and waveguide physical temperatures. In Table 5, the normalized T_{op} is computed from

¹ A. Freiley, *DSS 15 System Noise Temperature*, JPL Interoffice Memorandum (internal document), Jet Propulsion Laboratory, Pasadena, California, July 15, 1990.

$$T_{op,n} = T_{op} + (T'_{cb}/L_{wg})(1/L_{atm,s} - 1/L_{atm}) + (1/L_{wg})(T_{atm,s} - T_{atm}) + (T_{wg,s} - T_{wg}) \quad (2)$$

where T_{op} , T_{atm} , T_{wg} , and L_{atm} are, respectively, the average measured T_{op} and computed T_{atm} , T_{wg} , and L_{atm} values given in Table 5.

Values for standard DSS 13 atmospheric conditions at 8.45 GHz per DSN Document 810-5² are

$$T_{atm,s} = 2.17 \text{ K}$$

$$L_{atm,s} = 1.00814 \text{ (corresponding to 0.0352 dB)}$$

Other X-band (8.45 GHz) values used for Eq. (2) are

$$T'_{cb} = 2.5 \text{ K}$$

$$L_{wg} = 1.0163 \text{ (corresponding to 0.07 dB)}$$

$$T_{wg,s} = 4.69 \text{ K for the above } L_{wg} \text{ and a standard physical waveguide temperature of 20 deg C}$$

At X-band and higher microwave frequencies, weather changes can cause significant variations of atmospheric noise contributions to operating system noise temperatures. As may be seen in Table 5, the applications of weather corrections enabled significant improvements to be made in the comparisons of operating noise temperatures for the different test configurations even when the measurements were performed on different days (months) of the different test configurations. Figure 7 reveals the sensitivity of DSS 13 X-band atmospheric noise contributions as functions of typical weather parameters during a calendar year. It is of interest to note that, at X-band, different weather conditions during the year at DSS 13 can cause the atmospheric noise temperatures to vary between 1.9 K and 4.2 K. The peak-to-peak variation of 2.3 K for X-band (8.45 GHz) is small compared to a peak-to-peak variation of 27 K that was computed for Ka-band (32 GHz) for the same weather limits shown in Fig. 7.

A computer program, SDSATM7M.BAS,³ was used to derive values for Fig. 7. In this program the Earth's atmosphere is assumed to be 30 km high and is divided up into 300 layers, each of which is 0.1 km thick. The input data consists of the height of DSS 13 above sea level, barometric pressure, ground-level air temperature, and ground-level

relative humidity. From the input data, the average pressure, water vapor, and oxygen content are determined for each layer. In Method 1 of the program, a constant mean physical temperature of 265 K is assumed for each layer for the oxygen noise contribution. For the water vapor, every layer is assumed to be at a constant physical temperature that is 10 K lower than the measured ground-level physical temperature. The noise temperature is computed for each layer and then integrated to arrive at the total atmospheric noise temperature. A more rigorous Method 2 by Otoshi was incorporated into the program as an option to take into account the temperature gradient of water vapor and oxygen content in each layer. Comparisons of the results from the two methods revealed that Method 1, used for the results described in this article, is sufficiently accurate for making corrections for atmospheric noise temperature changes due to weather. It should be pointed out that Method 1 is also being used for making weather corrections to the BWG antenna efficiency measurement data.⁴

C. Tipping Curves

Tipping curve tests were performed at F1 and F3 at various azimuth angles for the test package. A few of the tipping curves are shown in Fig. 8. On a clear day, the tipping curves should be independent of azimuth angles except at elevation angles approaching those that correspond to the ground profile. A horizon profile map shows that the DSS 13 ground profile elevation angle is less than about 1 deg for 50- and 126-deg azimuth angles, but is about 7 deg at the 315-deg azimuth angle.

Due to the high levels of wide-angle side lobes on previous large Deep Space Network (DSN) antennas, it has not been possible to derive the atmospheric noise contributions from the measured operating system temperature values only at zenith and the 30-deg elevation angle. The general equation for extracting the zenith atmospheric noise temperature from tipping-curve data is derived as

$$T_{atmz} = L_{ant}[T_{op}(\theta_{el}) - T_{op}(90)]/[1/\sin(\theta_{el}) - 1] + T'_{cb}(1 - 1/L_{atmz}) \quad (3)$$

where

θ_{el} = the elevation angle, deg

$T_{op}(\theta_{el})$ = the operating system temperature measured at θ_{el} , K

² Deep Space Network Flight Project Interface Design Handbook, 810-5, rev. D., vol. I, module TCI-30 (internal document), Jet Propulsion Laboratory, Pasadena, California, June 1, 1990.

³ Courtesy of S. Slobin of the Jet Propulsion Laboratory, Pasadena, California. The SDSATM7M.BAS program is a modified version of SDSATM4.BAS, but gives the same answers.

⁴ S. Slobin, private communication, Jet Propulsion Laboratory, Pasadena, California, November 1990.

$T_{op}(90)$ = the operating system temperature measured at $\theta_{el} = 90$ deg or at zenith, K

L_{atmz} = the loss factor for the zenith atmosphere. An estimate of this value can be made by using the loss factor for a standard atmosphere, or a better estimate can be obtained from use of the SDSATM7M.BAS program for the actual weather conditions during the measurement.

T'_{cb} = the effective contribution to T_{op} from the cosmic background, K

L_{ant} = the loss factor from the antenna aperture to the input to the HEMT, K

For the special case of $\theta_{el} = 30$ deg, Eq. (3) becomes

$$T_{atmz} = L_{ant}[T_{op}(30) - T_{op}(90)] + T'_{cb}(1 - 1/L_{atmz}) \quad (4)$$

In practice, the factor L_{ant} is often set equal to 1.0 and the last term in Eq. (4) is often completely ignored. At 8.45 GHz and under standard atmospheric conditions, the last term contributes only 0.02 K. The values used in Eq. (4) to obtain the results in Table 6 were $T'_{cb} = 2.5$ K and $L_{ant} = L_{wg} = 1.0163$ (corresponding to 0.070 dB). For simplicity, the antenna reflector surface and BWG resistive losses were not included in the L_{ant} value.

Table 6 shows a comparison of the measured T_{atm} as obtained through the use of Eq. (4) and computed T_{atm} as obtained through the SDSATM7M program, which is based on the weather data shown. As seen in Table 6, the difference between measured and computed atmospheric noise temperatures is only ± 0.2 K at X-band for this antenna. The low amount of scattering from the slim tripod legs on the BWG antenna provides a possible explanation for this good agreement.

D. Subreflector Tests

Figure 9 shows a subreflector defocus curve obtained at F1 at an average elevation angle of 51 deg while tracking radio source 3C274. At the time tests were being done at F1, a good test procedure for obtaining subreflector defocus data with a radio source had not yet been fully developed. Calibrated and annotated strip-chart recordings were used to obtain required data. Hence, the F1 subreflector defocus data are estimated to be accurate only down to the -13-dB level. Also shown in Fig. 9 is a predicted Z-defocus curve that is based on a mechanical structural model by R. Levy and another curve based on a GTD/Jacobi-Bessel antenna program by P. Cramer. The Cramer calculations were done for a rigging angle of 45 deg

and, therefore, do not include any gravity effects.⁵ It can be seen in Fig. 9 that the agreements between predicts from the two theoretical models and the experimental data are quite good down to about -12 dB.

For verification of the performance of the antenna optics at F3, a subreflector Z-defocus test was performed in the region of the 45-deg elevation angle. In contrast to the strip chart method used for F1 tests, for F3 tests the operating system temperature data for each subreflector offset position were averaged by the computer and sent to both a line printer and a disk file. Figure 10 shows a comparison of predicts for F1 and experimental results at F3. It should be pointed out that even though Cramer's predicts were made for F1 and the experimental results are for F3, there is good agreement between theory and experiment. Note also that good agreement is still obtained between Cramer's predicts and experimental results at defocus positions in the side-lobe regions of the subreflector defocus curve. For a properly designed and aligned BWG mirror system, the subreflector defocus gain curves at F3 should be nearly identical to those at F1. Therefore, a Z-defocus subreflector test at F3 provides a good method for verifying the integrity of the BWG system.

A Z-defocus subreflector test is also useful for revealing whether multipath signals exist within a large antenna system [3,4]. Scrutiny of the subreflector defocus data near the peak of the main beam (Fig. 11) indicates there are no unusual humps, as were found on a 64-m S-band (2295 MHz) Cassegrain antenna [3]. The absence of humps and ripples on the main beam defocus curve shown in Fig. 11 indicates that the magnitudes of any multipath signals within the BWG system are probably small.

Unlike Y-defocus tests, when Z-defocus tests are performed, new sets of boresight readings do not have to be obtained at each new subreflector setting. However, a sufficient number of off-boresight and on-source readings need to be obtained to enable corrections to be made for potential receiver gain changes. A Z-defocus test can be performed rather quickly (1.5 hr). If too much time is taken to perform a subreflector test, the antenna characteristics (which are functions of elevation angle) will change sufficiently as to invalidate the test at a desired average elevation angle.

In order to obtain accurate subreflector defocus data below the 13-dB level (Figs. 9 and 10) while tracking a

⁵ P. Cramer, *Calculated DSS-13 Subreflector Z-axis Focus Curves, Feed at F1*, JPL Interoffice Memorandum 3328-90-0355 (internal document), Jet Propulsion Laboratory, Pasadena, California, September 20, 1990.

10 K source, the radiometer needs to provide resolutions and accuracies better than 0.2 K. In addition, the antenna-pointing model should be sufficiently good so as to allow blind pointing and remain on the peak of the source to within ± 4 mdeg.

E. Overall Performance

The overall performance of the X-band test package is summarized in Table 7. The performance data are based on observed test data in a worst-case field environment as well as on error analysis equations given in [5]. Examination of Table 7 reveals that the X-band test package performance was excellent in terms of resolution, gain stability, and linearity.

The test package was subjected not only to mechanical stresses during radio source tracking at various elevation

angles at F1, but also to ambient temperatures ranging from 0 to 40 deg C and transportation to and from the ground, F1, and F3 several times. Test data in Table 3 show that the long-term stability and repeatability of data for the various test configurations were very good.

IV. Conclusions

The goal of obtaining accurate noise-temperature values with the test package at the various focal points has been achieved. The X-band test package has operated well in all of the different test configurations and performed better than expected in terms of short- and long-term gain stability. To the authors' knowledge, the results presented in this article demonstrate the first known experimental method for determining the degradations caused by the BWG system of any large antenna.

Acknowledgments

Numerous persons have contributed to the acquisition of high-quality data. Many who contributed to the test package development have been previously acknowledged in Part I. Those who contributed to obtaining the results reported in Part II are acknowledged here. The Goldstone crew, supervised by J. Carpenter and J. May, installed the test package at F1 and F3 and their efforts are gratefully acknowledged. J. Garnica and G. Bury of DSS 13 assisted in obtaining the F1 noise-temperature diagnostics data given in Table 2. Scheduling conflicts for F1 and F3 tests were resolved by the Test Coordinator, M. Britcliffe of the Ground Antennas and Facilities Engineering Section. The continual on-site support of DSS 13 personnel and DSS 13 supervisor C. Goodson is gratefully acknowledged. Without the contributions and cooperation of both named and unnamed individuals in various groups throughout the Telecommunications Division and the assistance of Goldstone personnel, this project could not have been brought to a successful conclusion.

References

- [1] T. Y. Otoshi, S. R. Stewart, and M. M. Franco, "A Portable X-band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part I: Design and Ground Tests," *TDA Progress Report 42-103*, vol. July–September 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 135–150, November 15, 1990.
- [2] C. T. Stelzried, *The Deep Space Network—Noise Temperature Concepts, Measurements, and Performance*, JPL Publication 82-33, Jet Propulsion Laboratory, Pasadena, California, pp. 2-1–2-3, September 15, 1982.

- [3] G. S. Levy, D. A. Bathker, A. C. Ludwig, D. E. Neff, and B. L. Seidel, "Lunar Range Radiation Patterns of a 210-ft Antenna at S-band," *IEEE Transactions on Antennas and Propagation*, vol. AP-15, no. 2, pp. 311-313, March 1967.
- [4] T. Y. Otoshi and D. L. Brunn, "Multipath Tests on 64-m Antennas Using the Viking Orbiter-1 and -2 Spacecraft as Far-Field Illuminators," *DSN Progress Report 42-31*, vol. November and December 1975, Jet Propulsion Laboratory, Pasadena, California, pp. 41-49, February 15, 1976.
- [5] C. T. Stelzried, "Operating Noise-Temperature Calibrations of Low-Noise Receiving Systems," *Microwave Journal*, vol. 14, no. 6, pp. 41-46 and 48, June 1971.

Table 1. Symbols and abbreviations

Symbol	Definition
T_{cb}'	Effective noise temperature contribution to T_{op} from the cosmic background radiation, K. This value is a function of frequency and will differ from the actual cosmic background noise temperature of 2.7 K. (See [2] and Note at the end of this table.)
T_{cb}	Cosmic background radiation noise temperature (nominally 2.7 K)
h	Planck constant
f	Frequency, Hz
k	Boltzmann constant
T_{atm}	Atmospheric noise temperature, K
T_{wg}	Noise temperature due to waveguide loss between the horn aperture and the input flange of the HEMT, K
T_{hemt}	Effective noise temperature of the HEMT as defined at the input flange of the HEMT, K
T_{fup}	Effective noise temperature of the follow-up receiver (downconverter + cables + power meter, etc.) as defined at the input flange of the HEMT, K
T_{op}	Operating system noise temperature as defined at the input flange of the HEMT, K
T_s	Source noise temperature, K
RH	Relative humidity
HEMT	High-electron-mobility transistor
L_{wg}	Loss factor for waveguide between the horn aperture and the input flange of the HEMT, ratio
L_{atm}	Loss factor of the atmosphere, ratio

Note : $T_{cb}' = T_{cb} \left[\frac{x}{\exp(x) - 1} \right]$ where $x = \frac{hf}{kT_{cb}}$.

Table 2. Preliminary diagnostic tests at F1: sources of contributions to operating SNT at 8.45 GHz

No.	Contributor description	Noise temperature, K
1	Hoist panels removed	0.4
2	Hatch door open	0.1
3	Open areas at bases of tripod legs	0.2
4	Primary shroud opening for the bypass mode	0.4
5	Openings at the base of bypass shroud on dish surface	0.4
6	Open area below the feed horn for the center pass mode (for tests at F1 only)	0.6

Table 3. Summary of X-band zenith operating system temperatures at DSS 13, from June 10, 1990 to February 2, 1991

Configuration	Observation dates	Grand average ^a T_{op} , K	Peak deviations from grand average, K
Ground	06/10/90, 01/21/91, 01/26/91	22.7	+0.3 -0.3
F1	10/04/90	25.9	—
F3	11/06/90, 11/09/90	34.2	+0.1 -0.2
After mirrors and ellipsoid realigned on December 18, 1990			
F3	01/31/91, 02/02/91	34.8	+0.1 -0.1

^a See Table 5 for the average T_{op} for each observation period. These values formed the basis for obtaining the grand average for a particular test configuration.

Table 4. Differential zenith operating system temperatures for various test configurations at 8.45 GHz

Configurations differenced ^a	Delta T_{op} , K
F1-Ground	3.2
F3-F1	8.3
After mirrors and ellipsoid realigned on December 18, 1990	
F3-F1	8.9

^a See Table 3 for ground, F1, and F3 values.

Table 5. X-band (8.45 GHz) measured zenith operating system temperatures corrected for weather and waveguide loss changes

Configuration	Observation period	Average measured T_{op} , K	Average weather during observation	Computed T_{atm} , K	Computed L_{atm}	Physical waveguide temp., deg C	T_{wg} , K	Normalized ^a T_{op} , K
X-band, ground	06/10/90 DOY 161 0800-1200 UT	23.10	891.8 mbar 17.8 deg C 74.4% RH	2.89	1.0108	17.6	4.65	22.4
X-band, ground	01/21/91 DOY 021 1649-1652 UT	22.39	898.2 mbar 4.9 deg C 22.6% RH	2.16	1.0081	8.6	4.51	22.6
X-band, ground	01/26/91 DOY 026 0600-1200 UT	22.73	896.5 mbar 6.5 deg C 36.7% RH	2.22	1.0084	2.2	4.40	23.0
X-band, F1	10/04/90 DOY 277 0800-1200 UT	26.03	900.0 mbar 25.7 deg C 20.4% RH	2.36	1.0088	19.8	4.68	25.9
X-band, F3	11/06/90 DOY 310 0600-1500 UT	33.98	890.0 mbar 8.9 deg C 27.4% RH	2.16	1.0082	17.5	4.65	34.0
X-band, F3	11/09/90 DOY 313 0900-1600 UT	34.45	902.4 mbar 16.0 deg C 29.4% RH	2.32	1.0088	20.7	4.70	34.3
After mirrors and ellipsoid realigned on 91 DOY 018								
X-band, F3	01/31/91 DOY 031 0900-1500 UT	34.63	902.9 mbar 4.7 deg C 13.7% RH	2.11	1.0080	20.7	4.70	34.7
X-band, F3	02/02/91 DOY 033 0730-1330 UT	34.91	899.5 mbar 6.0 deg C 18.8% RH	2.13	1.0081	21.9	4.72	34.9

^a Normalized T_{op} values were computed through the use of Eq. (2).

Table 6. Measured and computed zenith atmospheric noise temperatures at X-band (8.45 GHz)

Observation period	Average weather	Configuration	Azimuth angle, deg	Computed ^a T_{atm} , K	L_{atmz}	$T_{op}(30) - T_{op}(90)$, K	Measured ^b T_{atm} , K	Delta T_{atm} , K
07/09/90 DOY 190 1950-2050 UT	895.9 mbar 30.9 deg C 20.8% RH	F1	50.9	2.48	1.0092	2.30	2.36	0.12
07/09/90 DOY 190 1950-2050 UT	895.9 mbar 30.9 deg C 20.8% RH	F1	126.5	2.48	1.0092	2.41	2.47	0.01
10/25/90 DOY 298 1638-1706 UT	895.0 mbar 26.0 deg C 13.0% RH	F3	50.0	2.19	1.0083	2.16	2.22	-0.03
01/30/91 1991 DOY 30 2110-2143 UT	902.0 mbar 10.5 deg C 10.2% RH	F3	315.0	2.09	1.0080	2.27	2.33	-0.24

^a Computed values were obtained from the computer program SDSATM7M.BAS.
^b Measured value obtained from zenith and 30-deg elevation T_{op} values of the tipping curve, and Eq. (4).

Table 7. X-band test package performance characteristics

Parameter	Performance achieved
Receive polarization	RCP, LCP, or fixed linear
Receive frequencies	8.4-8.5 GHz (determined by fixed local oscillator and IF filter in downconverter)
T_{op} for test package on the ground at DSS 13	< 23 K
Gain stability over 0 to 40 deg C	< 0.3 dB p-p, < 0.05 dB/hr (ovenized heating temperature control)
Delta T_{op} resolution, $\tau = 1$ sec, 100 MHz bandwidth	< 0.02 K
Total calibration nonlinearity error	< 1%
T_s measurement accuracy	$\pm[0.02 + 0.010 \times T_s]$ K for $2 < T_s < 10$ K
Overall T_{op} measurement accuracy, K ($10 < T_{op} < 150$ K)	< 0.4 K p-p

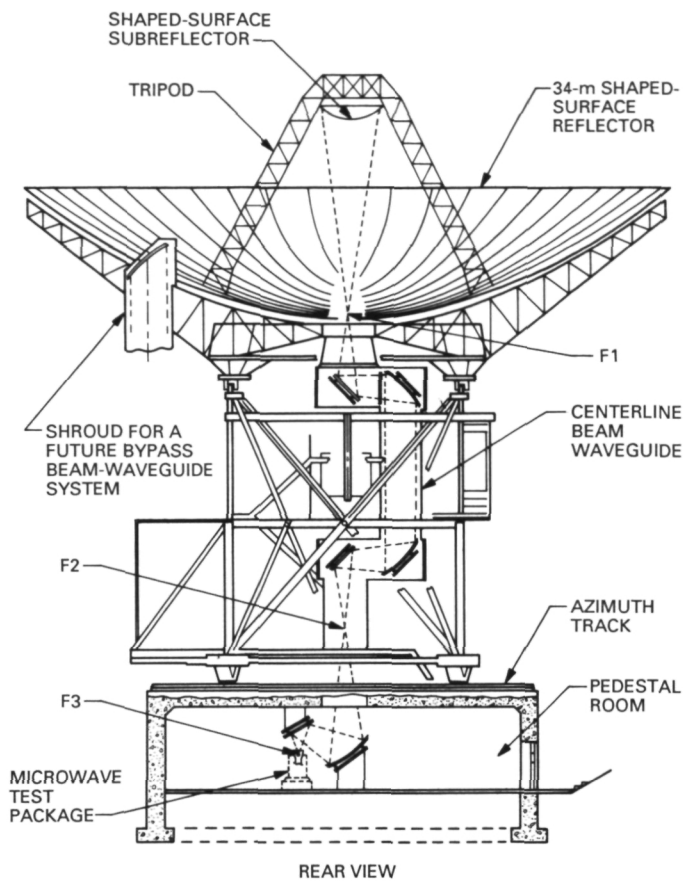


Fig. 1. Beam-waveguide antenna depicting focal points F1, F2, and F3.

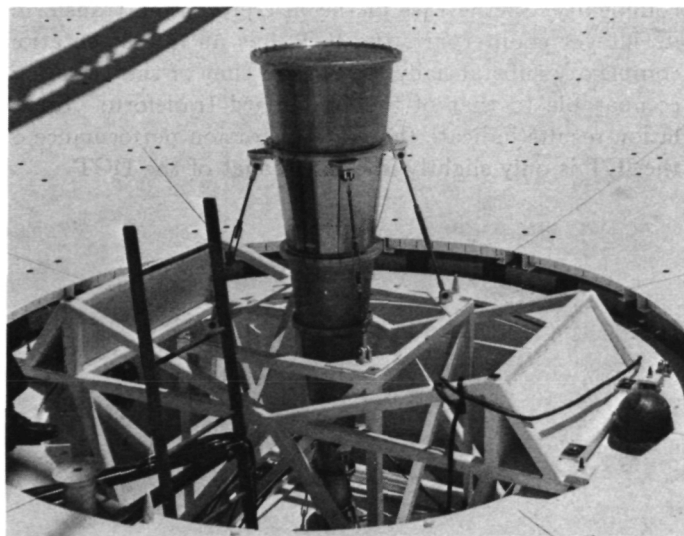


Fig. 2. X-band 29-dBi horn test package and mounting assembly installed at F1.

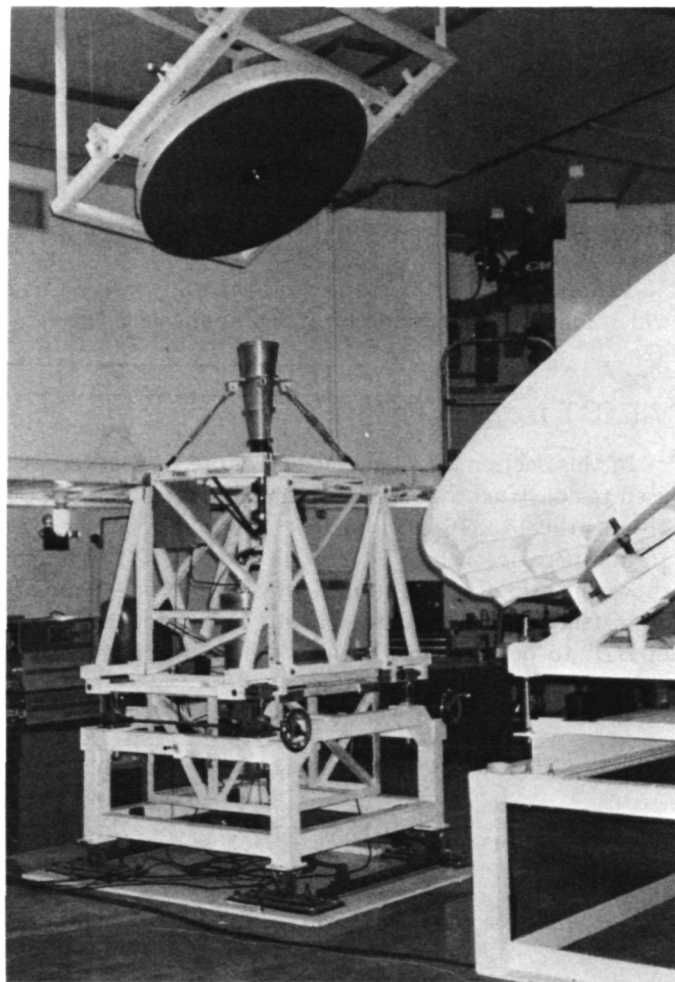


Fig. 3. X-band 22-dBi horn test package and mounting table installed at F3.

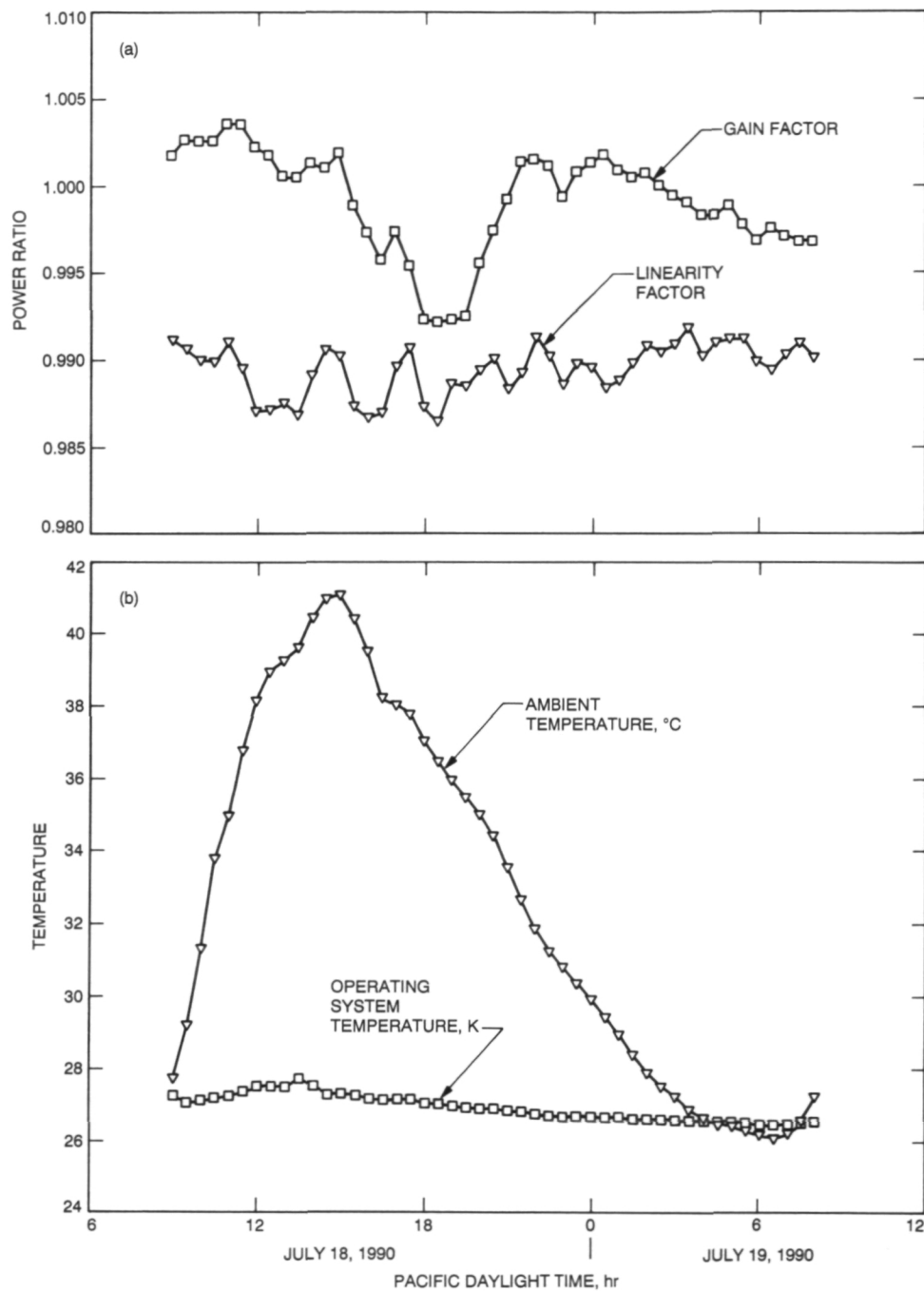


Fig. 4. Mini-cal data taken on July 18, 1990 with the X-band 29-dBi horn test package mounted at F1: (a) gain factor and linearity factor, and (b) operating system noise temperature and ambient temperature.

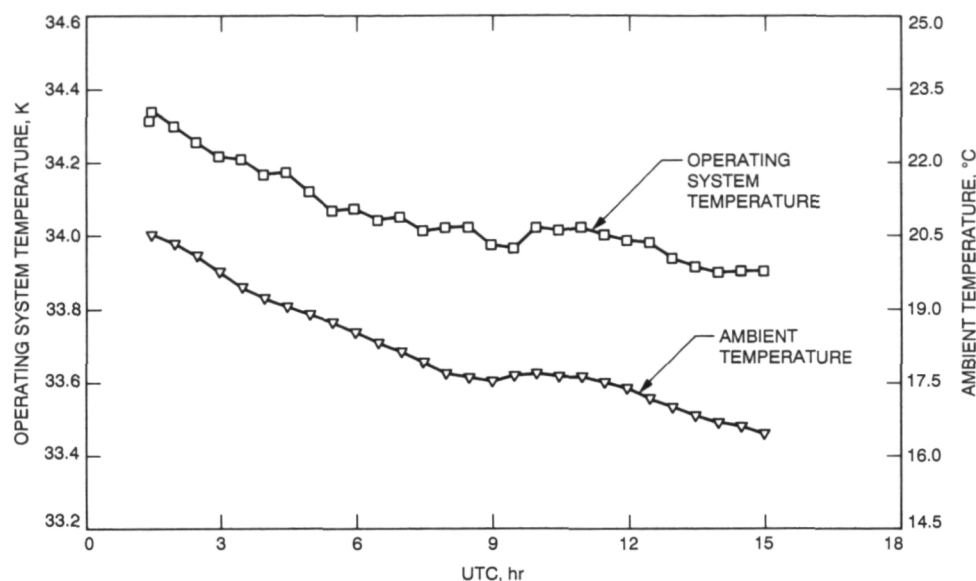


Fig. 5. Operating system noise-temperature mini-cal data taken on November 6, 1990 with the X-band 22-dBi horn test package mounted at F3.

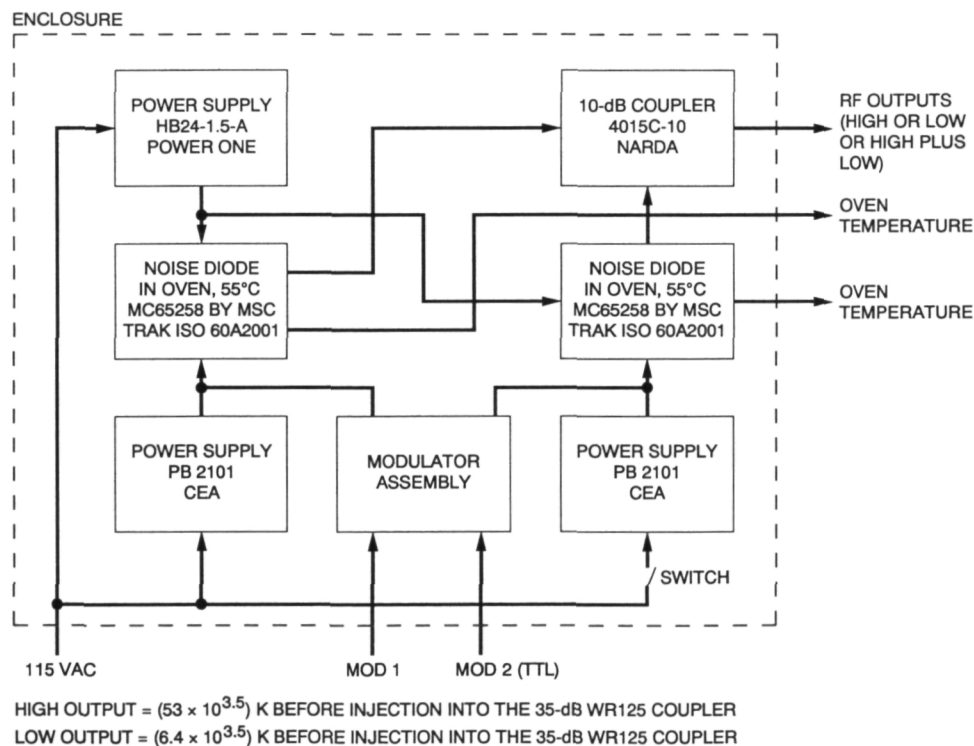


Fig. 6. X-band noise box assembly (mounted inside the X-band test package).

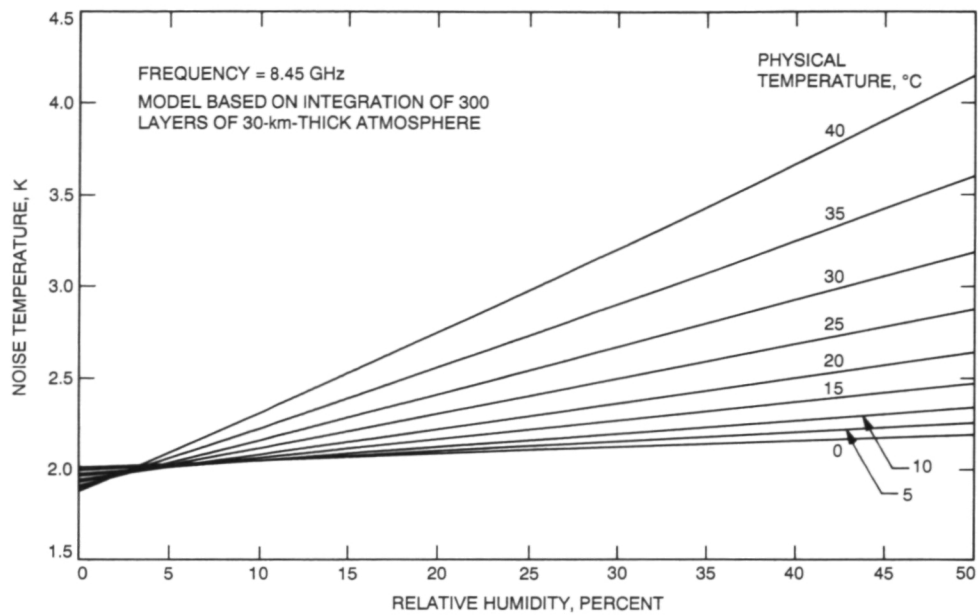


Fig. 7. Theoretical X-band atmospheric noise-temperature contributions versus weather parameters at DSS 13 (courtesy of S. Slobin).

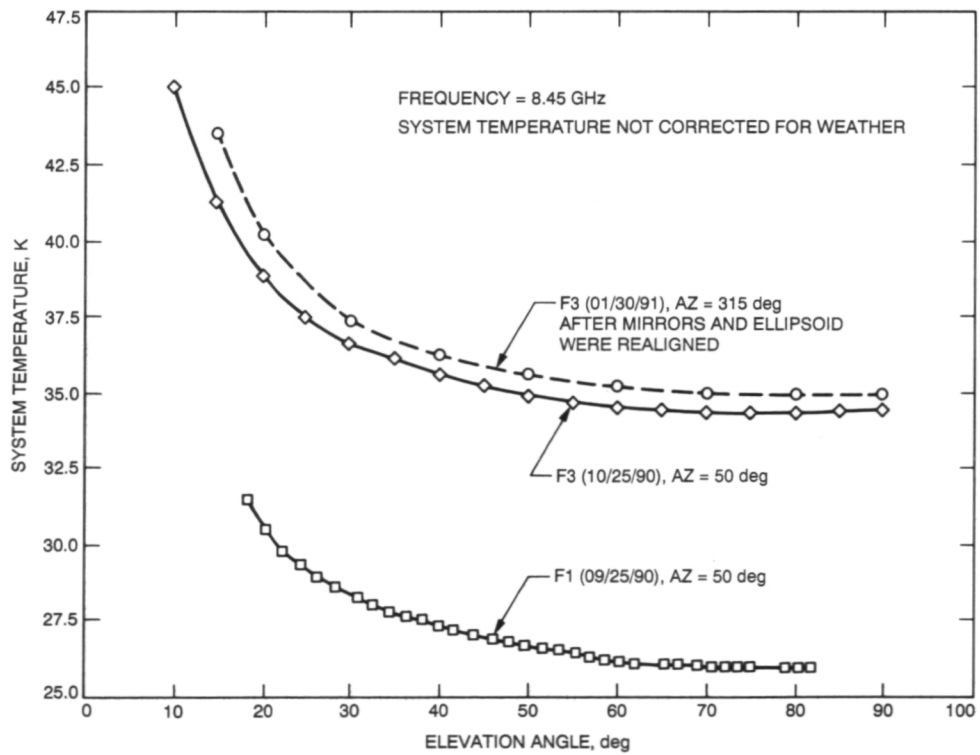


Fig. 8. X-band tipping curves at F1 and F3.

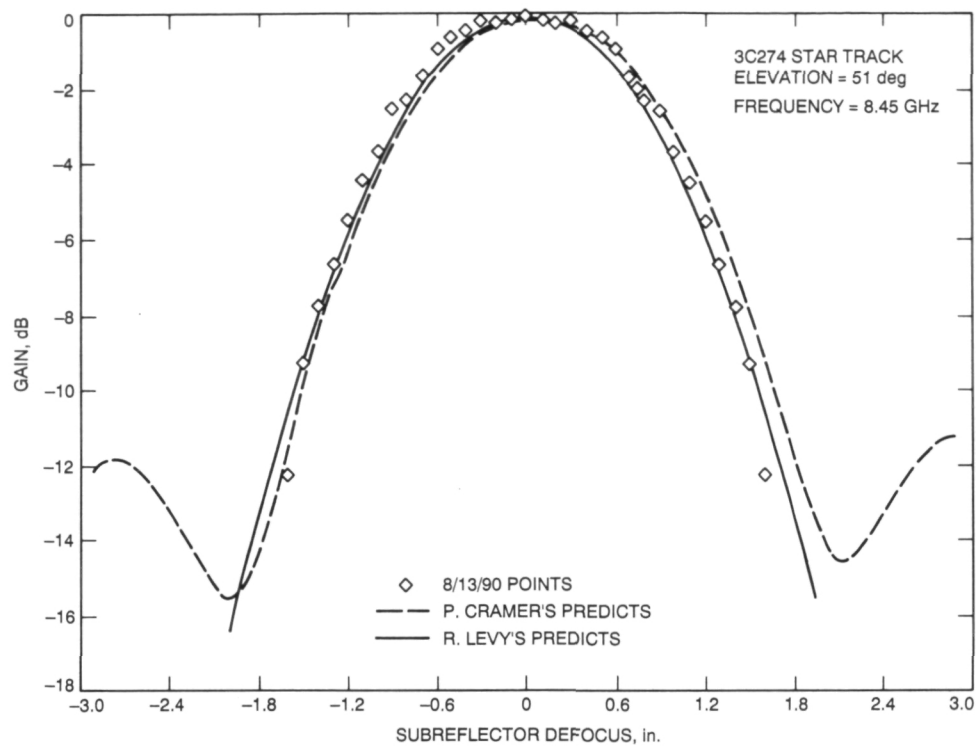


Fig. 9. X-band subreflector Z-defocus curve measured at F1.

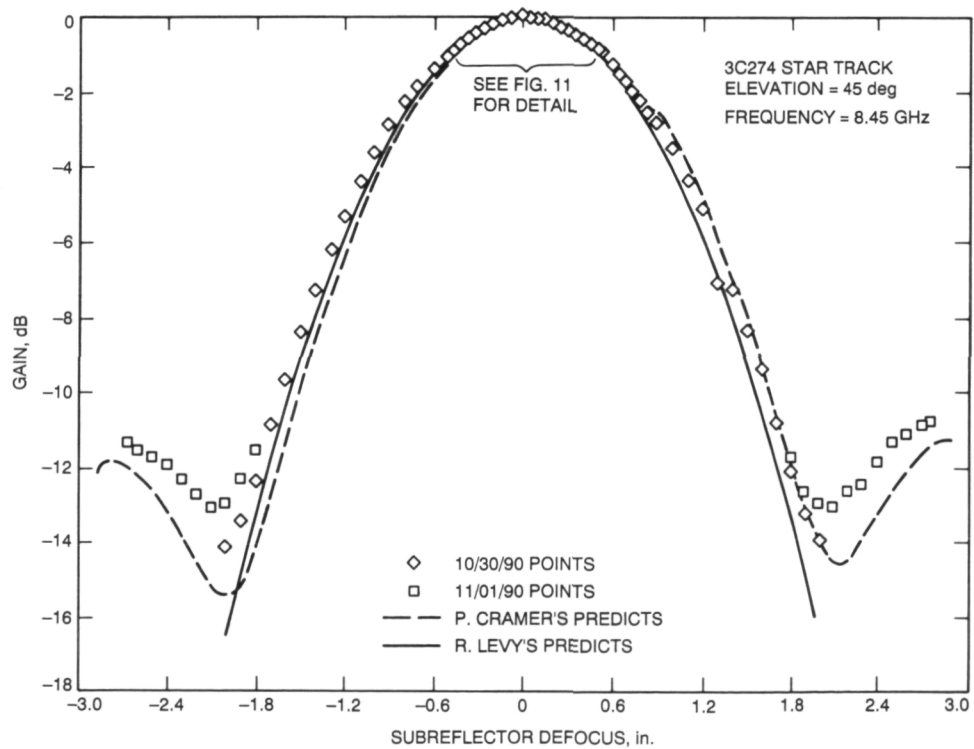


Fig. 10. X-band subreflector Z-defocus curve measured at F3.

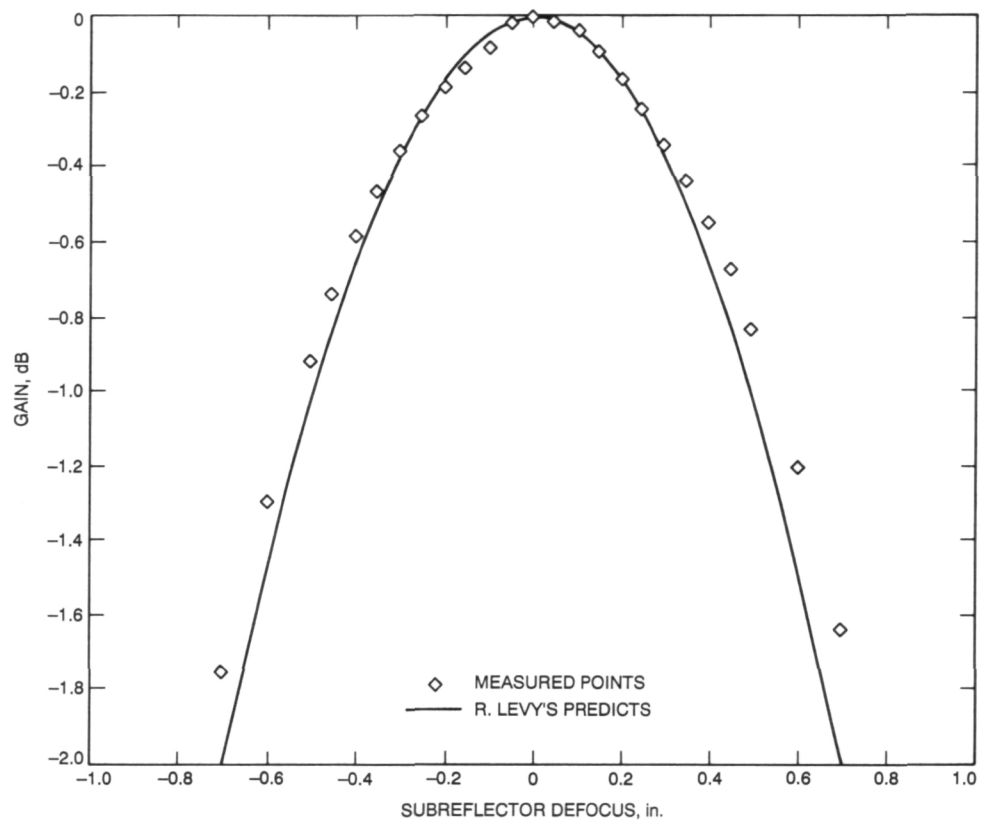


Fig. 11. Expanded detail of main beam of the F3 X-band subreflector defocus curve of Fig. 10.

1994025118

N94- 29621

TDA Progress Report 42-105

May 15, 1991

442609

Phasing the Very Large Array on Galileo in the Presence of Jupiter's Strong Radio Emission

J. S. Ulvestad
Navigation Systems Section

Work is in progress to determine the feasibility of using the Very Large Array (VLA) radio telescope to receive telemetry from Galileo during its close encounter with Io on December 7, 1995. The VLA was used previously to receive telemetry from Voyager 2 at Neptune. However, Jupiter's strong radio emission is an additional complication in the case of the Galileo encounter. This article analyzes the effect of Jupiter's radio emission on the phase-adjustment procedure ("autophasing") used to maintain coherence among the 27 VLA antennas. Results of an experiment designed to mimic the Io encounter are presented. As expected, Jupiter's strong radio emission has a considerable effect on the autophasing procedure. A simple emission model is found to give a good approximation to the fringe-visibility plots derived from the VLA data, and that successful model is used to estimate the VLA's ability to autophase on Galileo during the Io encounter. The effect of Jupiter should be small for projected baselines longer than ~ 800 m, and completely negligible for projected baselines longer than ~ 1.1 km.

The most extended configuration of the VLA (the **A** configuration) probably can be used successfully for telemetry reception during the Io encounter. Further analysis and testing of the effect of correlated noise from Jupiter is necessary before a final decision can be made about the feasibility of using the second largest (**B**) configuration of the VLA for reception of Galileo telemetry. Use of the **B** configuration could simplify the upgrades needed to support the Io encounter. Tests to help choose the preferred VLA configuration could be performed by using the VLA to observe the Magellan spacecraft at Venus during July and October 1991.

Examination of the effects of planet noise on the VLA have implications beyond the use of that telescope for supporting the Io encounter. The effects of planet radio emission on spacecraft data received by antenna arrays are relevant to choosing the exact locations of antennas that might be built by the Deep Space Network in coming years.

I. Introduction

The Galileo encounter with Io, scheduled for December 7, 1995, requires reception of 8.4-GHz (X-band) telemetry at 134.4 kilobits/sec for approximately 4 hours. The Galileo Array Study Team was formed in the spring of 1990 to determine how the Deep Space Network (DSN) could provide sufficient confidence in the data return from the Io encounter. The report of that team was given in [1]. An important conclusion was that use of the Very Large Array (VLA) radio telescope [2,3] would provide a powerful enhancement to the DSN reception capability for the Io encounter. The VLA already has been used successfully by the DSN, as it was arrayed with the Goldstone Deep Space Communications Complex to return telemetry from the Voyager 2 encounter with Neptune [4]. The expected use for the Io encounter would differ in that the VLA would be used in a stand-alone mode, rather than having its signal arrayed with that of the Goldstone Complex.

The VLA, located southwest of Socorro, New Mexico, consists of 27 25-m antennas, nine on each arm of a Y-shaped configuration. These antennas are ~60 percent efficient at 8.4 GHz and are equipped with receivers having zenith system temperatures of ~30 K [5]. The spacings between adjacent antennas along each arm increase with distance from the center of the array. Four major configurations are employed: **A**, **B**, **C**, and **D**. The respective total lengths of each arm in these configurations are 21 km, 6 km, 2 km, and 0.6 km.

The VLA would use the spatially unresolved signal from the Galileo spacecraft to adjust the phase calibration of each antenna in order to maintain coherence among the different antennas of the array. A detailed investigation of this procedure for the Voyager 2 Neptune encounter was reported in [6]. The "autophasing" procedure relies upon the assumption that the received signal on each interferometer baseline is dominated by a point source, so that phase residuals are dominated by system noise and propagation effects, rather than by extended radio emission. More discussion of the meaning of autophasing and the phase residuals is found below, in Section II.

The presence of the strong radio emission of Jupiter places constraints on the ground configuration. This planet is an extended radio source with a flux density of ~8 to 20 Jy¹ at 8.4 GHz; the exact strength depends on the variable Earth-Jupiter distance. In an observing bandwidth of ~8 MHz, which was the value used for the Voy-

ager encounter with Neptune, Jupiter's signal will be considerably stronger than that expected from Galileo when the spacecraft arrives at the planet. Therefore, autophasing can be corrupted by the presence of Jupiter. However, Jupiter is an extended radio source; as baseline lengths are increased, the spatial resolution of Jupiter's emission reduces Jupiter's effect on the autophasing process. In addition to its effect on autophasing, Jupiter's radio emission can affect the signal-to-noise ratio (SNR) of the telemetry stream obtained by adding the signals of the individual antennas. The effect of Jupiter will be to cause some correlation between the noise at antennas that are located close to each other. This correlated noise would be more harmful than uncorrelated noise; analysis of its effects is beyond the scope of this article.

On September 28, 1990, Jupiter passed approximately 2 arcmin from the natural radio source 0839+187. This radio source has an 8.4-GHz flux density near 1 Jy, similar to the expected Jupiter-encounter value for Galileo in an 8-MHz bandwidth; also, like Galileo and other spacecraft, it is an unresolved source as seen by the VLA.

This article presents an analysis of the effect of Jupiter on VLA autophasing on the Galileo spacecraft. Section II defines interferometer phases and elaborates on the autophasing process. Section III contains a calculation of the relative strengths expected for Jupiter and Galileo on the day of Io encounter. Section IV analyzes the theoretical effect of the resolution of Jupiter on autophasing using the VLA. Section V describes VLA observations of Jupiter's close passage to 0839+187 on September 28, 1990; a model that adequately reproduced the fringe visibility during that passage is described. In Section VI, that model is used to predict the effects expected on the day of Io encounter in 1995. Section VII contains a discussion of the results of the simulations and their impact on the selection of the VLA configuration for the Io encounter. Possible further VLA experiments are described in Section VIII. Broader implications for the locations of antennas in future DSN antenna arrays are discussed briefly in Section IX. Finally, the results of this work are summarized in Section X.

II. The Meaning of Phases and Autophasing

In order to understand the VLA phase-adjustment process and the experiment described below, it is necessary to describe the meaning of "phase" and "autophasing" in interferometric observations. A two-element interferometer measures the amplitude and relative phase of the cross-correlation of a signal for a pair of antennas; the VLA

¹ 1 Jy = 10^{-26} W m⁻² Hz⁻¹

is a set of 351 simultaneously operating two-element interferometers. The measured phase on each baseline is affected by the geometry of the observations, propagation effects, instrumental effects, and system noise. Both the propagation effects and many of the instrumental effects are antenna-based in nature; i.e., they are peculiar to a particular antenna and common to all two-element interferometers including that antenna. Therefore, standard calibration methods use measured interferometer phases to determine phase offsets peculiar to each antenna. The normal interferometer mode of the VLA involves no attempt to adjust phases in real time, since those phases carry information about the structure of a radio source that is being mapped. Instead, calibration is accomplished by making observations of another (point) radio source, which are used later to solve for the phase calibration of the individual antennas. During post-observation data analysis, that calibration is applied to the source that is to be mapped.

Autophasing differs from the interferometer mode of the VLA because it attempts to make all the antennas coherent on a particular source by means of a quasi-real-time phase calibration. This is done for telemetry reception and for very long baseline interferometry in order to make the VLA mimic a single antenna whose collecting area is equivalent to that of a 130-m-diameter radio telescope. The calibration is made during the observations by assuming that the source being observed is a point source. Then the phase contributed by the source structure should be negligible. The interferometer phases on each baseline are used to construct a least-squares solution that finds the phase calibration peculiar to each antenna, and then remove the antenna-dependent effect from the data in (near) real time. Phases are corrected by means of a feedback loop delayed by a few seconds to tens of seconds. Both the 351 measured baseline phases and the 27 individual antenna phase calibrations will be referred to below.

III. Relative Strengths of Jupiter and Galileo

At the time of the Io encounter, the distance from the Earth to Galileo and Jupiter will be 6.2 astronomical units (AU), or 9.3×10^8 km. The total flux density of Jupiter at 8.4 GHz should be approximately 8.4 Jy (scaled from values given in [7]). About 6.7 Jy would be in the planet's atmospheric emission, with another 1.7 Jy in the radiation belts. A 25-m VLA antenna has a sensitivity of 0.110 K/Jy at 8.4 GHz [5]. Therefore, the antenna temperature contributed by Jupiter will be about 0.9 K on the day of Io encounter. This is a small fraction of the system temperature of an individual antenna; thus, Jupiter should add

only a small amount to the random thermal noise at each antenna.

For Voyager, an 8-MHz observing bandwidth was used for autophasing. In that bandwidth, the total power received from Jupiter during the Io encounter would be $6.7 \times 10^{-19} \text{ W m}^{-2}$. However, the centimetric radio emission of Jupiter is unpolarized, whereas the spacecraft signal is all in a single circular polarization. Thus, in comparing the spacecraft strength to Jupiter, only a single circularly polarized VLA channel should be considered; the received power from Jupiter in such a channel will be only half the total power. Taking this factor into account and considering the scaling with the observing bandwidth $\Delta\nu$, the received power in a single VLA channel will be

$$P_J = 3.4 \times 10^{-19} \left(\frac{\Delta\nu}{8 \text{ MHz}} \right) \text{ W m}^{-2} \quad (1)$$

For Galileo, estimates for the low-power mode of the high-gain antenna are a transmitter power of 40.5 dBm with a mean gain of 48.7 dBi in the direction of Earth. Spacecraft power is expected to be inadequate for use of the high-power mode of the transmitter, so the low-power mode is assumed throughout this article. At a distance of 6.2 AU, the expected received power is

$$P_G = 7.7 \times 10^{-20} \text{ W m}^{-2} \quad (2)$$

This value assumes that the total observing bandwidth used at the VLA will be wide enough to contain all significant power from various harmonics of the Galileo telemetry sidebands. Therefore,

$$\frac{P_J}{P_G} = 4.4 \left(\frac{\Delta\nu}{8 \text{ MHz}} \right) \quad (3)$$

Equation 3 holds for $\Delta\nu \gtrsim 5 \text{ MHz}$, which is wide enough to encompass almost all the power in the sidebands; at smaller bandwidths, the received power in the Galileo signal would begin to be reduced.

At this stage, it should be noted that VLA observations of Galileo probably will use a different setup than for the Voyager-Neptune encounter. Specifically, rather than operating the VLA correlator in its continuum mode for the autophasing, the correlator will be used in the spectral-line mode. The different setup is related to the need to eliminate significant telemetry gaps in VLA stand-alone data for a data rate that will be more than six times the rate

delivered from Voyager at the Neptune encounter. Consideration of the details of the setup of the back end of the VLA are beyond the scope of this article, but those details will influence the final choice of the effective autophasing bandwidth, $\Delta\nu$.

IV. Predicted Effect of the VLA's Spatial Resolution of Jupiter

At VLA rise time on the day of the Io encounter, Galileo and Jupiter will be separated by 0.85 arcmin, well within the primary beam of a single VLA antenna, which has a full-width at half maximum of ~ 5 arcmin at 8.4 GHz. If Jupiter were a point source, its signal would be considerably stronger than that from Galileo (cf. Section III above); the VLA autophasing procedure would make the antennas mutually coherent on Jupiter, but then they would not be in phase for Galileo. Therefore, the Galileo telemetry signal from the 27 antennas would not be combined coherently, and the effective gain of the VLA would not give an adequate SNR for the spacecraft telemetry.

Interferometers having a long enough baseline can resolve away the bulk of Jupiter's disk, so that the autophasing procedure would respond primarily to the point-like radio emission from Galileo. To evaluate the baseline length required, consider the complex fringe visibility for an interferometer with baseline \vec{b} . This interferometer observes an extended radio source whose intensity distribution is $I(\vec{r})$, where \vec{r} is the angular position in the radio source relative to a reference point within the source. The amplitude of that fringe visibility, multiplied by the total power from the source, gives the correlated power observed by the interferometer. The complex fringe visibility is given by

$$V = \frac{1}{A_0 I_0} \int_S A(\vec{r}) I(\vec{r}) e^{-2\pi i \vec{b} \cdot \vec{r} / \lambda} d\Omega \quad (4)$$

(e.g., [8]). In Eq. (4), $A(\vec{r})$ is the effective area of an individual antenna in the direction given by \vec{r} , A_0 is the effective antenna area along the pointing direction, I_0 is the total intensity of the radiation from Jupiter, λ is the observing wavelength, S is the source area that contributes to the integral over the solid angle Ω , and i is the imaginary number representing $\sqrt{-1}$.

Assume that the VLA antennas' effective areas are constant over the radio source, since the individual antenna beamwidths are much larger than Jupiter's disk. Further,

let the distribution of emission be uniform over the Jovian disk (angular radius r_J) and zero elsewhere. Then $A(\vec{r})I(\vec{r}) = A_0 I_0$ for $r \leq r_J$. [The assumption that $I(\vec{r})$ is constant may lead to an overestimate of Jupiter's effect on the autophasing by up to 20 percent, since it ignores the fact that some of the radio emission comes from more extended radiation belts.] Then

$$V \approx \int_0^{r_J} r dr \int_0^{2\pi} e^{-2\pi i \vec{b} \cdot \vec{r} / \lambda} d\theta \quad (5)$$

where θ is the angle between \vec{r} and the projection of \vec{b} on the sky plane. If D is the projected baseline length as seen from the direction of Jupiter, Eq. (5) reduces to

$$V \approx \int_0^{r_J} r dr \int_0^{2\pi} e^{-2\pi i D r \cos \theta / \lambda} d\theta \quad (6)$$

Note that D is the *projected* baseline perpendicular to the line of sight to Jupiter; the projected baseline, rather than the actual physical separation of two antennas, is the relevant quantity for interferometry. The integral in Eq. (6) gives a first-order Bessel function, J_1 . Defining $D_\lambda \equiv D/\lambda$ to be the projected baseline length in units of wavelengths, the effective power from Jupiter would be

$$P_{\text{eff},J} = P_J |V| = P_J \frac{|J_1(2\pi D_\lambda r_J)|}{\pi D_\lambda r_J} \quad (7)$$

Suppose the effective power from Jupiter were less than 20 percent of the power received from Galileo on a particular VLA interferometer. Then Jupiter would affect the phase on that baseline by less than 0.2 rad. Summing two antennas out of phase by such a small angle gives a very small amplitude loss for the Galileo telemetry (cf. [6]). Hence, a reasonable condition for successful autophasing is

$$\frac{P_G}{P_{\text{eff},J}} \geq 5 \quad (8)$$

Combining Eqs. (3), (7), and (8) leads to the following condition for successful autophasing:

$$\frac{|J_1(2\pi D_\lambda r_J)|}{\pi D_\lambda r_J} \frac{\Delta\nu}{8 \text{ MHz}} \leq 0.046 \quad (9)$$

This equation holds as long as $\Delta\nu$ is not so small that a significant fraction of the Galileo signal power is excluded

from the autophasing bandwidth. Examination of the values of the Bessel function (e.g., [9]) shows that, assuming $\Delta\nu = 8$ MHz, the requirement given by Eq. (9) is met for $2\pi D_\lambda r_J \geq 10$ (r_J in radians). On December 7, 1995, the angular radius of Jupiter will be 15.2 arcsec (74 μ rad). This implies that, for 8.4-GHz observations, the projected interferometer spacing should be larger than 0.77 km. Many of the projected baselines in the VLA C configuration, used for the Neptune encounter, are shorter than this. Projected baselines will tend to be shortest when Galileo rises at the VLA, since the spacecraft clears the 8-deg elevation limit at an azimuth similar to that of the southeast arm of the VLA. Thus, the analysis shows that the VLA must be arranged in the larger B or A configuration for autophasing to be successful at the Io encounter.

V. Test Observations of Jupiter's Close Passage to 0839+187

A. The Observations

On September 28, 1990, at 1000 Universal Time (UT), Jupiter passed less than 2.1 arcmin from the natural radio source 0839+187. This natural source is a primary calibrator at the VLA, unresolved on all baselines at 8.4 GHz. Since Jupiter rose at the VLA at 0920 UT, there was an opportunity to use this close passage to test the effect of Jupiter on the ability of the VLA to autophase on the nearby point radio source. The distance to Jupiter was approximately 6 AU, making its flux density and angular size slightly larger than those expected for the Galileo encounter with Io. Approximately 70 minutes of VLA time were scheduled on September 28, from 0955 to 1105 UT. At that time, the VLA was in a B/C hybrid configuration (north arm in the B configuration, others in the C configuration). Projected antenna spacings ranged from about 70 m to 6.7 km (2,000 to 190,000 wavelengths).

During the tests, a point-source calibrator (0851+202) about 3 deg from the field of interest was observed in normal interferometer mode (with no phase feedback loop) and in autophasing mode in order to set an approximate flux density scale and to check the system performance. (See Section II above for a further description of these observing methods.) The field including 0839+187 and Jupiter also was observed in both modes in order to supply data for autophasing simulations and to provide real-time autophasing data. Finally, an autophasing observation was made with the VLA split into two subarrays, one containing only antennas with mutual baselines longer than 850 m and the other containing many antennas near

the central hub of the VLA, with a mix of long and short baselines. This last observation was meant to test the real-time performance of the autophasing procedure when only the longer antenna spacings were included.

Figure 1 shows an image of 0839+187 and Jupiter, calibrated using the observations of the nearby calibrator 0851+202. This image has undergone a deconvolution process to remove the effects of sidelobes in the synthesized VLA beam. The apparently elongated shape of the point source is caused by asymmetry in the response function of the VLA due to the north arm of the VLA being much longer than the other two arms during these observations. The *peak* flux density of the point source was a factor of ~ 7 higher than that of Jupiter because of the resolution of Jupiter with the longest VLA baselines. However, the *total* flux density of Jupiter was much higher than that of the point source.

B. Test Results

The results of the observations of 0839+187 initially were confusing. The SNR of more than 15 to 1 in a 5-sec integration on 0839+187 would have given phase residuals on the order of 4 deg on each baseline if system noise were the limiting factor. During observations in the autophasing mode, the displayed residual phases for each *antenna* indicated that the individual antennas had small phase residuals, as though the autophasing were working well. However, the phases on many individual *baselines* were quite large and remained relatively constant, sometimes in the vicinity of 60 deg. A listing of the phases on all baselines after the experiment indicated high values in many cases. In particular, interferometer phases on baselines that included any one of three antennas that were moderately close to the center of the array were on the order of 90 deg. Baselines involving only the few antennas closer to the center and those farther out had phases that were less than 20 deg.

The real-time antenna phase listing was misleading because the "residuals" displayed actually were just the differences between the antenna phases at the current integration and those predicted based on past integrations. These residuals accurately depict the system performance when the dominant error consists of dynamic phase fluctuations due to the troposphere or to other causes. However, in the situation where there is extended radio emission, there are phase "errors" caused by the erroneous assumption that the radio source being observed is a point source. In that case, the antenna phase solutions may not fluctuate greatly from one integration to the next even though phase coherence is not being maintained on the radio source of

interest. The feedback loop would not generate phase coherence on the point source, even though the lack of fluctuations would cause the reported residuals to be small.

In the post-observation data processing, attempts were made to solve for the complex antenna gains (amplitude and phase) from the normal interferometry data. This procedure is identical to the real-time process used in autophasing, but without a feedback loop. The solutions gave enormous errors; convergence sometimes did not occur. (Unfortunately, similar failures of convergence are not reported in real time during autophasing.) Often, no individual phase solutions were reported for the same three antennas that had shown the large baseline phases in the post-experiment listings of the autophasing data. This may have been due to the confusion caused by the mixture of short and long baselines involving these antennas, with the short baselines dominated by Jupiter and the long baselines dominated by 0839+187. Figure 2 shows a portion of the reported errors from the least-squares program that attempts to compute antenna gains; this output is from an individual 5-sec integration made in normal interferometry mode at 1025 UT. The huge amplitude and phase closure errors show that autophasing would not have worked adequately.

Off-line gain solutions also were computed for the normal interferometry data using various limitations on the baseline lengths. Elimination of all baselines whose projections were shorter than 20,000 wavelengths (0.71 km) gave decent results, since the longer baselines detected very little correlated flux from Jupiter. Figure 3 displays reported errors (with closure limits of 5 percent in amplitude and 5 deg in phase) for the gain solutions with this minimum baseline length, using the same 5-sec integration as in Fig. 2. Although the baseline limitation eliminated some of the 351 baselines from the phase solution, the returned solution gave individual phases for each of the 27 antennas, since even the antennas near the center of the VLA had projected baselines longer than the specified cutoff when combined with the outermost antennas. As long as the solution for each antenna is not limited by a poor SNR on the individual baselines, the antenna solutions derived from a subset of all possible baselines should be adequate. For Galileo at Jupiter, the total received power on an individual baseline will be ~ 15 times higher than it was for Voyager 2 at Neptune, so an adequate SNR for phase determination will be achievable without using all the VLA baselines.

When the VLA was split into two subarrays performing autophasing independently of each other, the subarray containing only baselines longer than 850 m (most well

over a kilometer) performed quite well. No phase residuals larger than 15 deg were observed either in the real-time autophasing or in the listings generated in the post-processing of the data. Again, this validated the prediction that projected baselines longer than ~ 0.7 to 0.8 km (20,000 wavelengths) would be adequate to remove most of the effects of Jupiter.

C. Simulations of the Test Observations

A useful tool in understanding the robustness of autophasing is a plot of the amplitude of the fringe visibility *versus* baseline length (a "visibility plot"). If such a plot is made for interferometer observations when Jupiter is in the field of view, baselines short enough to see a significant correlated amplitude from Jupiter also would have their phases affected by the Jupiter emission. Thus, the effects of Jupiter on autophasing can be estimated by looking at visibility plots.

Figure 4 is a visibility plot derived from a 9-minute observation of 0839+187 and Jupiter, using the normal interferometer mode. On baselines shorter than about 35,000 wavelengths (1.25 km), the effects of Jupiter are noticeable; the effect is greater than ~ 20 percent for baselines shorter than about 20,000 km, as expected. Oscillations in the correlated amplitude are related to alternating constructive and destructive interference between Jupiter and the point source, and die out as baselines become long enough to resolve Jupiter completely. For baselines longer than 35,000 wavelengths, the flatness of the plot is a result of the absence of significant correlated flux from Jupiter. These baselines "see" only the point source 0839+187, indicating that autophasing would perform well in this regime.

Figure 5 shows a visibility plot for a simulation of the observation considered in the preceding paragraph. A simple model was used for the brightness distribution of the point source and Jupiter. The size of Jupiter's disk and the source separation were fixed by Jupiter's actual angular diameter and the positions of the point source and the planet, while the baselines sampled were those actually used to acquire the test data illustrated in Fig. 4. The flux densities of the sources in the simulation were adjusted to give a good qualitative correspondence between Figs. 4 and 5. The final model included a point source of flux density 0.93 Jy and a uniform disk of total flux density 6.0 Jy.² The absolute flux densities in the model are more

² Jupiter was actually somewhat stronger than 6.0 Jy, but its effective flux density was reduced because of the 2.5-arcmin half-power half-widths of the VLA antenna beams at 8.4 GHz.

uncertain than their ratio, since the amplitude calibration was based on a highly variable radio source whose strength was known only to ~ 10 percent accuracy. For each point, 0.054 Jy of noise (computed noise level for the parameters used in the test observations) was added with random phase.

Figure 6 is a comparison of the plots from Figs. 4 and 5, and shows the good correspondence between the real data and the simulation. Differences can be accounted for by the crude method by which noise was added in the simulation, the fact that different subsets of the real projected baselines were selected for the two plots, and the over-simplification of the uniform disk model for Jupiter.

VI. Simulations of the Io Encounter

The good correspondence between the real and simulated visibility data in the 1990 test indicated that a simple emission model could serve well as a predictor of the visibility function to be expected from the combined emission of Galileo and Jupiter on encounter day, December 7, 1995. The 1990 test was performed with Jupiter at a distance of 6 AU, giving it an angular size similar to that at the time of the Io encounter, when it will be 6.2 AU from Earth. Further, the model value of 6.5 for the ratio of the flux density of Jupiter to that of 0839+187 on September 28, 1990 was somewhat larger than the expected ratio of 4.4 for the Jupiter/Galileo ratio on the day of the 1995 encounter [cf. Eq. (3)], assuming an 8-MHz bandwidth. Thus, the test results give a reasonably conservative approximation of the expectations for the Io encounter.

On December 7, 1995, when Galileo and Jupiter rise at the VLA, a little less than 3 hours prior to Galileo's closest approach to Io, Jupiter will be located about 51 arcsec from Galileo. Figure 7 shows the predicted visibility plot for the model of Galileo and Jupiter, assuming use of the B configuration of the VLA, in which each of the three VLA arms has a total length of ~ 6 km. This plot was derived in the same way as the simulation displayed in Fig. 5, which was shown to be a fairly good representation of reality. More than a third of the baselines will have projected lengths shorter than 30,000 wavelengths and show some effect due to Jupiter. About 15 to 20 percent of the baselines, those shorter than $\sim 20,000$ wavelengths, show an effect greater than 20 percent. The shortness of the projected baselines is partly due to the fact that the spacecraft and Jupiter will be at -23 deg declination at the encounter. Therefore, they rise nearly along the

azimuth of the southeast arm of the VLA, giving considerable foreshortening of many baselines. If Jupiter's center and Galileo were located at the same point, Fig. 7 would look like a superposition of a constant and a Bessel function $J_1(x)/x$ (where x is proportional to the baseline length), the respective Fourier transforms of a delta function and a uniform disk. In fact, the plot is somewhat more complicated because Galileo will not be directly in front of Jupiter, so that Jupiter and Galileo will be beating against each other as Jupiter is being resolved.

Figure 8 shows the visibility plot for the same model as for Fig. 7, with the exception that the VLA antennas are assumed to be in their most extended arrangement, the A configuration. The baselines are more than three times longer than those in Fig. 7, and less than 10 percent of the baselines would be short enough to be affected significantly by Jupiter.

Figure 9 is a plot similar to Fig. 7 for the B configuration, except the assumed VLA bandwidth is 4.7 MHz instead of 8 MHz. The narrower bandwidth corresponds to the situation in which six spectral-line channels of 781 kHz each, centered on the Galileo carrier, are used for the phase adjustment process. Since Galileo's signal is restricted in frequency, whereas Jupiter is a broadband emitter, this bandwidth gives a higher SNR for Galileo and a lower SNR for Jupiter. The effective bandwidth cannot be made much smaller than ~ 5 MHz without reducing the total power from Galileo.

Using the 4.7-MHz bandwidth in Eq. (9) and evaluating the Bessel function, it appears that Jupiter should cause an effect of 20 percent or more only for baselines shorter than about 10,000 wavelengths (350 m). Figure 9 confirms the analytic approximation; about 8 percent of the baselines will be shorter than the critical length. Figure 10 is a plot made using the same assumptions for observations using the VLA A configuration. It shows a significant effect from Jupiter on fewer than 10 baselines.

Two hours after Galileo and Jupiter rise at the VLA, Galileo will have moved within 19 arcsec of Jupiter's center, appearing just off the limb of the planet. The spacecraft and Jupiter will be at 25 deg elevation. The projected baselines will be somewhat longer for the higher spacecraft elevation. Calculations show that, as expected, Jupiter's effects on the visibility function will be less pronounced even though it is closer to the spacecraft; this is a direct consequence of the increase in the projected baseline lengths. Thus, the limiting case in selection of the appropriate VLA configuration is the situation at the spacecraft rise time.

VII. Selecting the Proper VLA Configuration for the Io Encounter

The purpose of the VLA tests and simulations has been to select the appropriate parameters for reception of telemetry from Galileo at the Jupiter encounter. There are two issues to be considered. One is the autophasing of the VLA antennas, while the other is the problem of correlated noise from Jupiter in the summed signal from all the antennas. Problems with the autophasing can be reduced by setting a software limit on the minimum projected length for the baselines used in the autophasing procedure. However, even though this will enable the phase adjustment to converge on antenna solutions that maintain coherence of the signal from Galileo, the antennas close to the center of the array also may have small relative phase for some portion of the signal from Jupiter. Therefore, they will be affected by correlated noise from Jupiter that will degrade the SNR of the telemetry stream. Because of the correlated component of the noise, the effect of Jupiter on the system's sensitivity will be greater than predicted from the 0.9 K increase in system temperature at the individual antennas. Although the details of correlated noise are beyond the scope of this article, the correlated noise will be minimized when the effective Galileo/Jupiter power ratio is maximized. If no baselines show an effect in the visibility plot, there should be little effect from correlated noise. An ideal situation would result from a selection of parameters such that the visibility plots (see Figs. 7-10) would be dominated by Galileo and show little effect from Jupiter.

Consider the effect of Jupiter on the visibility function. Equation (8) specifies that Jupiter should contribute less than ~20 percent to the total amplitude of the fringe visibility. If Jupiter contributes less than that amount on a given baseline, its contribution to the autophasing error for that baseline will be only ~12 deg, and the correlated-noise contribution should be correspondingly small. As is evident from Fig. 7, assuming an 8-MHz bandwidth and observations using the **B** configuration, roughly 20 percent of the 351 baselines show too large an effect due to Jupiter. Therefore, this set of parameters may be inadequate for telemetry reception from Galileo. Figure 10 shows that for a 4.7-MHz bandwidth and the **A** configuration, only a few baselines will show bad effects due to Jupiter, so this set of parameters could be used for telemetry reception.

The remaining uncertainty is whether both the **A** configuration and the reduced bandwidth are necessary, or if only one change from the parameters of Fig. 7 would suffice. Inspection of Fig. 9 shows that, with a 4.7-MHz bandwidth and the **B** configuration, fewer than 10 percent of the baselines show a contribution of more than

20 percent due to Jupiter. For the **A** configuration and the 8-MHz bandwidth, Fig. 8 shows a 20 percent effect on about 5 percent of the baselines. These baselines can be eliminated easily in the autophasing solution, but the degradation of telemetry by Jupiter's correlated noise still must be evaluated.

The ideal choice would be to use the largest available configuration and smallest possible autophasing bandwidth for the Io encounter. However, making the reliability adequate for the **A** configuration would cost more than for the smaller **B** configuration, because of the decaying power-distribution system at the VLA; cable replacement would be needed to the end of the **A** configuration (21-km arms) instead of only to the end of the **B** configuration (6-km arms). Therefore, the feasibility of using the **B** configuration merits further study. Using the narrowest possible autophasing bandwidth would put more severe operational constraints on the spacecraft frequency predictions in order to prevent loss of the telemetry SNR due to the exclusion of significant sidebands.

The limiting factor for VLA observations of the Io encounter probably will be the correlated noise. The effect on autophasing could be reduced by software limits on the baseline lengths used in the autophasing solutions and/or by including a model of the Jovian emission in the autophasing process. However, the correlated noise could be more difficult to eliminate. At first glance, it would seem that the effect of the noise could be bounded by the increased noise expected on a single large antenna due to the system-temperature increase caused by Jupiter. However, the situation may be worse. A single large antenna samples all projected baselines included within the aperture, whereas an array of smaller antennas samples only a limited set of projected baselines. For example, a single 35-m antenna would have a system-temperature increase of only ~1.8 K due to Jupiter at a distance of 6 AU. But if the same aperture area is made up of two 25-m antennas, and the observing geometry is unfavorable, it is possible that destructive interference and correlated noise could cause the spacecraft signal to disappear completely, equivalent to an infinite increase in system temperature! Further investigation of the correlated noise should be conducted to choose the best configuration for the Io encounter.

VIII. Possible Future VLA Tests

Tests of autophasing in the larger VLA configurations, and particularly tests of the effects of correlated noise, should be conducted. Either a natural radio source containing both point and extended components, or a close

passage of Jupiter to another bright, compact radio source might be used to evaluate the noise in the summed signal of all the antennas. However, use of a natural radio source would not give a particularly accurate representation of the distribution of emission from Galileo and Jupiter. The best "near miss" between Jupiter and a strong point source (P 1352-104) will not occur until late 1993, when the VLA is scheduled to be in the **A** configuration. Even then, the angular separation between Jupiter and the compact radio source will be nearly twice as large as the separation in the 1990 test, nearly as large as the half-power beamwidth of the individual antennas at 8.4 GHz.

Observation of a spacecraft with discrete telemetry sidebands could be more useful than observing a natural radio source with a wideband noise-like signal. However, use of a spacecraft probably would provide additional benefits only if the telemetry signal were extracted and the signal properties (SNR, bit error rate, etc.) were studied. Without telemetry extraction, a test involving spacecraft observations would be similar to a test involving observations of a natural, pointlike radio source.

An ongoing test opportunity exists involving the Magellan spacecraft, which transmits an 8.4-GHz signal to Earth from its orbit around Venus. Appropriate selection of observing parameters probably could be made to mimic the situation when Galileo encounters Jupiter. Parameters that need to be considered include the VLA configuration, the Magellan declination, the Earth-Sun-Venus angle, the Venus flux density, the apparent angular size of Venus, and the spacecraft transmission times compared to its rise times at the VLA. Depending on phase angle, Venus has a brightness temperature of ~ 600 K at 3.6 cm [10], between three and four times that of Jupiter. At an Earth-Venus distance of 8.3×10^7 km, Venus has the same angular size (~ 15 arcsec) that Jupiter will have when Galileo arrives in 1995; therefore, it would have three to four times Jupiter's flux density. At this distance, the *total* Earth-received power from Magellan will be approximately 200 times that expected from Galileo on December 7, 1995. However, it may be possible to mimic Galileo's Io encounter by selecting the appropriate passband for observations of Magellan, so that only a small fraction of the total power is detected.

In July and October 1991, Venus will have the appropriate angular size of 15 arcsec and the VLA will be in the **A** (July) and **B** (October) configurations. Although Venus will not be at the right declination to get the same projected baselines as for Galileo's Io encounter, this still would be the best time to do further autophasing tests and investigations of the correlated noise problem. Results of such tests can enable a final decision to be made about

the feasibility of using the **B** configuration for Galileo's Jupiter encounter.

IX. General Considerations for DSN Antenna Siting

This article has addressed the issue of the proper VLA configuration for the Io encounter at some length. However, the work done also has a significant impact on the general question of the siting of new antennas for the DSN. Future deep-space telecommunications may rely more heavily on arrays of smaller antennas than on single large antennas such as the 70-m antennas currently in operation at each DSN complex. Indeed, arraying of a variety of antennas for telemetry reception has been used for 20 years; the chronology of that arraying was summarized in Table 1 of [4].

Inevitably, future spacecraft supported by the DSN will visit planets that are strong radio emitters at the planned telecommunication frequencies of 8.4 GHz and 32 to 34 GHz. The general trend for those spacecraft will be for them to involve orbiters that may spend years in the vicinity of a single planet. The Cassini spacecraft, which will arrive at Saturn a few years after the turn of the century, is an example of a candidate for support from a new generation of antenna arrays. It is essential that the geometry of those arrays take into consideration the effect of planetary radio noise, which can hinder the coherent phasing of the antennas, the reception of telemetry, and the adequacy of the SNR on radiometric data such as Doppler tracking. If the antenna sites are selected primarily for operational convenience, they may not be used to full advantage for spacecraft in planetary orbits. Antenna separations should be large enough that the effects of planetary radio emission can be reduced to insignificance by appropriate observing and data analysis techniques.

X. Summary

An analysis has been performed to determine the ability of the VLA to autophase on the Galileo spacecraft in the presence of the strong 8.4-GHz radio emission from Jupiter. This autophasing process must be successful in order for signals from the individual VLA antennas to be added coherently. The total power from Jupiter is expected to be about 4.4 times that from Galileo in a single, 8-MHz-bandwidth, circularly polarized channel. Calculations show that projected interferometer baselines longer than ~ 0.8 km will be necessary in order to resolve Jupiter adequately so that the autophasing is successful.

A VLA test has been performed that simulated the Galileo-Jupiter encounter that will occur in late 1995. This test showed that the autophasing process was corrupted severely when antennas were spaced closely enough to see a large amount of correlated emission from Jupiter. The corruption became relatively insignificant for projected interferometer baselines longer than 800 meters, in good agreement with the predictions. A simple model of the structure of the radio emission gave good success in reproducing the observed fringe visibility.

A simple emission model has been used to predict the fringe visibility and the effects of Jupiter at the Galileo encounter. Results show that the **A** configuration of the VLA should be adequate to resolve Jupiter enough to minimize both autophasing problems and correlated noise in

the telemetry stream. The **B** configuration may be adequate for autophasing if the effective bandwidth is reduced to ~ 5 MHz, but further analysis must be done to investigate the effects of the correlated noise. Observations of the Magellan spacecraft in orbit around Venus could provide important information about the feasibility of use of the VLA **B** configuration for data return from Galileo at Jupiter. Such tests could be done in July and October of 1991.

The tests using the VLA may have a broader applicability for the DSN. The problems of receiving telemetry and radiometric data from planetary orbiters in the presence of radio-emitting planets are relevant to the selection of sites for individual antennas in future arrays that might be built by the DSN.

Acknowledgments

The author thanks Roger Linfield for using his occultation-prediction software to identify the close passage between Jupiter and 0839+187 and for searching for similar events with Venus, Jupiter, and Saturn in the 1991 to 1995 period. The author also thanks Rachel Dewey and Don Brown for carefully reading and commenting on a draft of this article.

References

- [1] J. W. Layland, P. E. Beyer, D. W. Brown, R. W. Burt, R. C. Clauss, J. M. Ludwinski, B. D. Madsen, J. C. McKinney, F. D. McLaughlin, D. J. Mudgway, N. A. Renzetti, J. S. Ulvestad, and R. J. Wallace, "Galileo Array Study Team Report," *TDA Progress Report 42-103*, vol. July-September 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 161-169, November 15, 1990.
- [2] A. R. Thompson, B. G. Clark, C. M. Wade, and P. J. Napier, "The Very Large Array," *Astrophysical Journal Supplement Series*, vol. 44, pp. 151-167, October 1980.
- [3] P. J. Napier, A. R. Thompson, and R. D. Ekers, "The Very Large Array: Design and Performance of a Modern Synthesis Radio Telescope," *Proceedings of the IEEE*, vol. 71, pp. 1295-1320, November 1983.
- [4] D. W. Brown, W. D. Brundage, J. S. Ulvestad, S. S. Kent, and K. P. Bartos, "Interagency Telemetry Arraying for Voyager Neptune Encounter," *TDA Progress Report 42-102*, vol. April-June 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 91-118, August 15, 1990.

- [5] J. S. Ulvestad, G. M. Resch, and W. D. Brundage, "X-band System Performance of the Very Large Array," *TDA Progress Report 42-92*, vol. October–December 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 123–137, February 15, 1988.
- [6] J. S. Ulvestad, "Phasing the Antennas of the Very Large Array for Reception of Telemetry from Voyager 2 at Neptune Encounter," *TDA Progress Report 42-94*, vol. April–June 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 257–273, August 15, 1988.
- [7] G. L. Berge and S. Gulkis, "Earth-Based Radio Observations of Jupiter: Millimeter to Meter Wavelengths," in *Jupiter*, edited by T. Gehrels, Tucson: University of Arizona Press, pp. 621–692, 1976.
- [8] A. R. Thompson, "The Interferometer in Practice," in *Synthesis Imaging in Radio Astronomy*, edited by R. A. Perley, F. R. Schwab, and A. H. Bridle, Astronomical Society of the Pacific Conference Series, vol. 6, pp. 11–37, 1989.
- [9] S. M. Selby, ed., *Standard Mathematical Tables*, 21st edition, Cleveland, Ohio: Chemical Rubber Company, p. 538, 1973.
- [10] W. M. Kaula, *An Introduction to Planetary Physics: The Terrestrial Planets*, New York: Wiley and Sons, 1968.

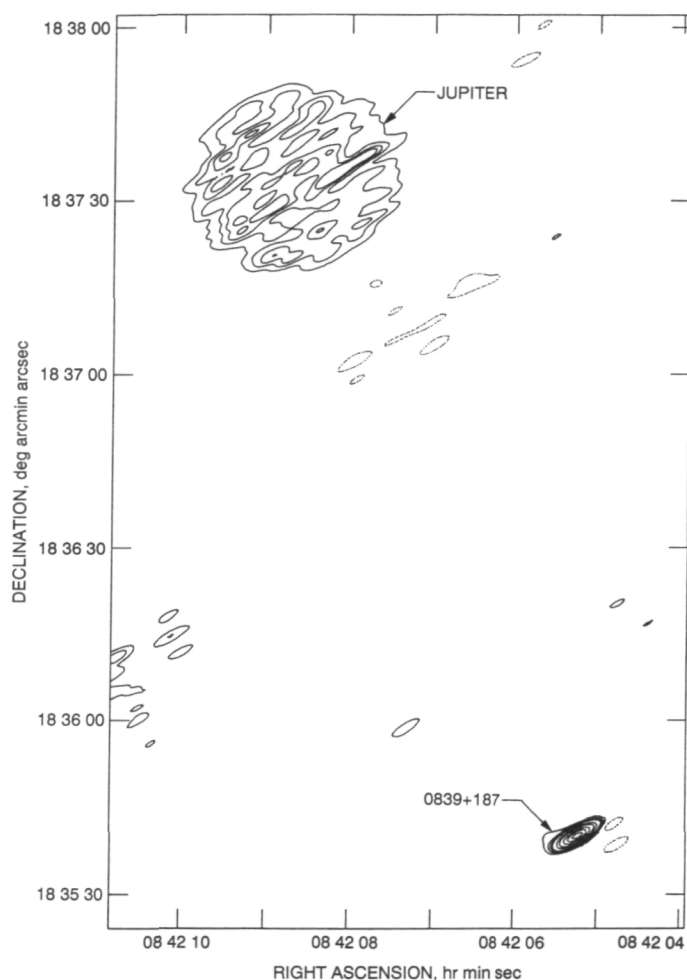


Fig. 1. X-band image of Jupiter and the radio source 0839+187, as taken by the VLA on September 28, 1990.

Closure errors at 0/10 25 0.0839+187									
1-2	184%	-3d	1-4	-47%	-9d	1-5	135%	158d	
1-6	163%	1d	1-7	220%	-9d	1-8	137%	2d	
1-9	84%	22d	1-10	158%	-31d	1-11	732%	4d	
1-12	327%	16d	1-13	165%	2d	1-14	167%	13d	
1-15	170%	-6d	1-16	114%	-17d	1-17	421%	-133d	
1-18	-35%	-22d	1-19	-48%	101d	1-20	606%	45d	
1-21	141%	11d	1-22	178%	-2d	1-23	162%	-7d	
1-24	-35%	-37d	1-25	113%	-11d	1-26	155%	-8d	
1-27	190%	-4d	1-28	-1%	85d	2-4	-35%	2d	
2-5	180%	165d	2-6	174%	3d	2-7	237%	-3d	
2-8	190%	-3d	2-9	88%	1d	2-10	169%	-1d	
2-11	129%	1d	2-12	212%	-5d	2-13	165%	-4d	
2-14	180%	10d	2-15	167%	2d	2-16	98%	-6d	
2-17	501%	-116d	2-18	-21%	-29d	2-19	-13%	92d	
2-20	234%	3d	2-21	184%	4d	2-22	198%	4d	
2-23	200%	2d	2-24	-31%	-32d	2-25	183%	-14d	
2-26	224%	14d	2-27	206%	4d	2-28	30%	85d	

Fig. 2. Reported amplitude (percent) and phase (d) errors from the least-squares solution attempting to compute VLA antenna gains, during a 5-sec integration on 0839+187 when it was near Jupiter. All baselines were used in the solution. Only errors involving antennas 1 and 2 are shown.

Closure errors at 0/10 25 0.0839+187								
1-7	13%	-5d	1-8	-11%	3d	1-21	0%	11d
2-16	-16%	6d	2-18	11%	0d	2-19	-11%	-3d
2-20	17%	-3d						

Fig. 3. Reported errors from the least-squares solution attempting to compute VLA antenna gains, during interferometry observations of 0839+187 when it was near Jupiter. Only baselines with projected lengths longer than 20,000 wavelengths were used. The original data were the same as the data used for Fig. 2. Errors involving antennas 1 and 2 are shown.

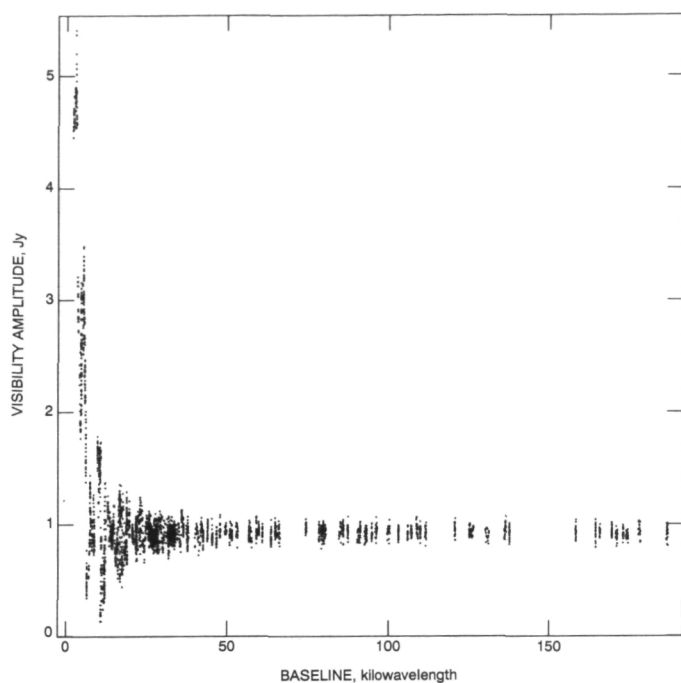


Fig. 4. Visibility plot for VLA observations of 0839+187 and Jupiter on September 28, 1990.

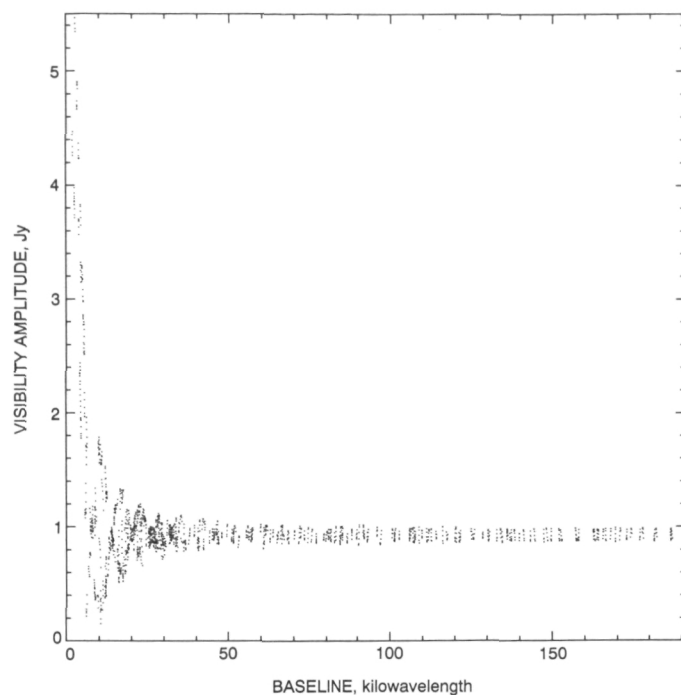


Fig. 5. Visibility plot computed from a simple emission model for the VLA test observations on September 28, 1990.

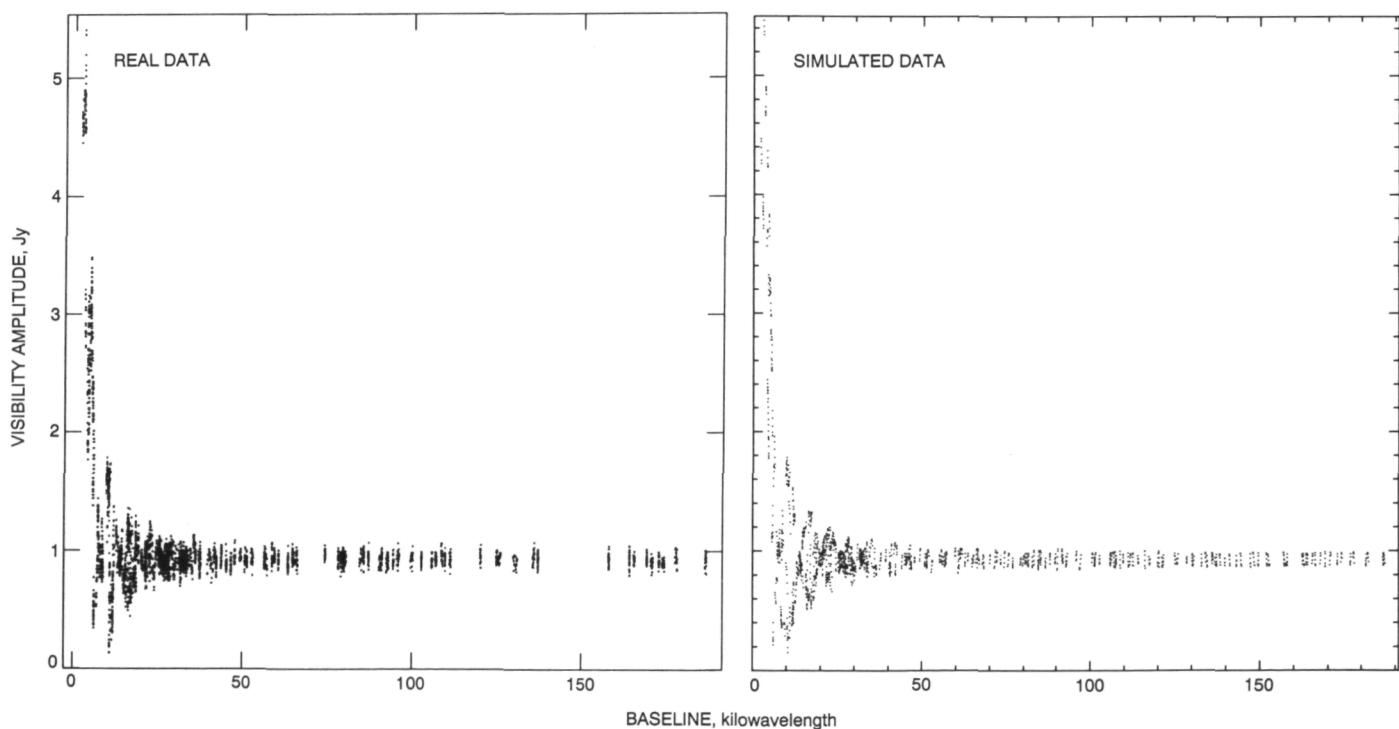


Fig. 6. Comparison of data (Fig. 4) and simulation (Fig. 5) for the VLA observations on September 28, 1990.

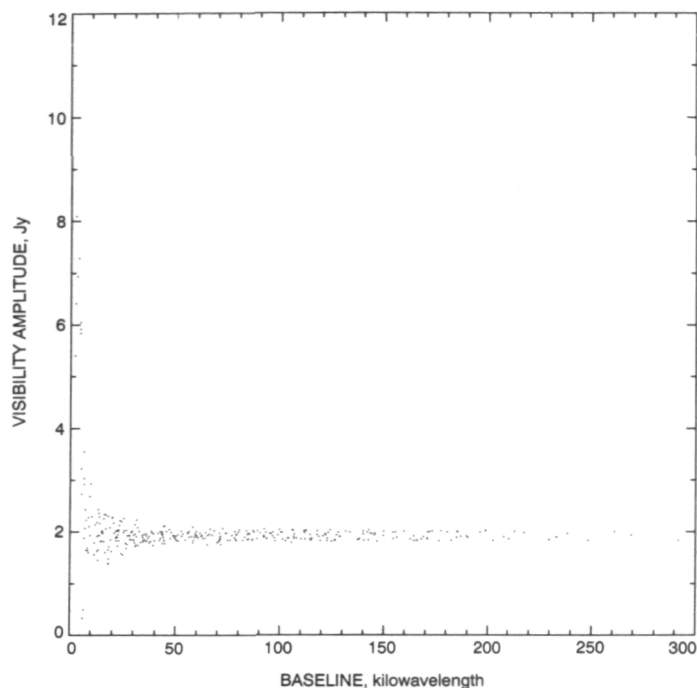


Fig. 7. Simulated visibility plot for observations of Galileo and Jupiter at their rise time (8 deg elevation) at the VLA on December 7, 1995. An 8-MHz bandwidth and use of the VLA B configuration are assumed.

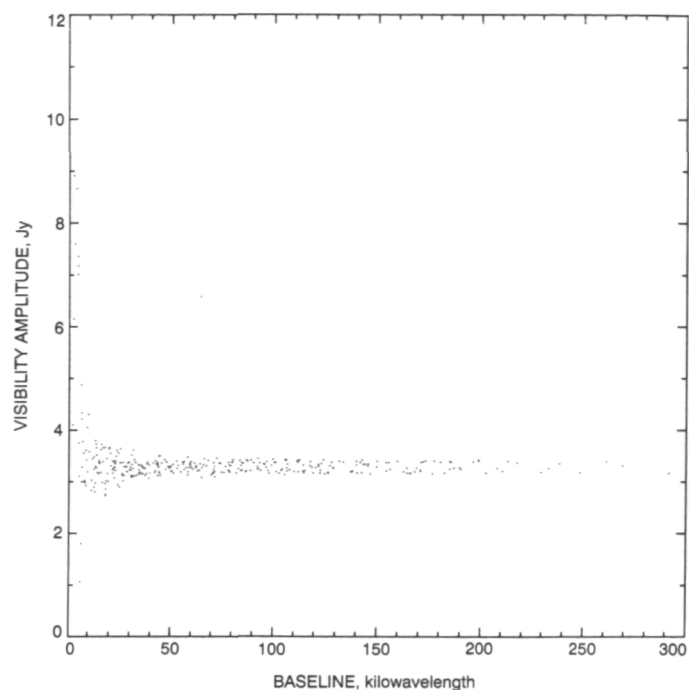


Fig. 9. Simulated visibility plot for observations of Galileo and Jupiter at their rise time (8 deg elevation) at the VLA on December 7, 1995. A 4.7-MHz bandwidth and use of the VLA B configuration are assumed.

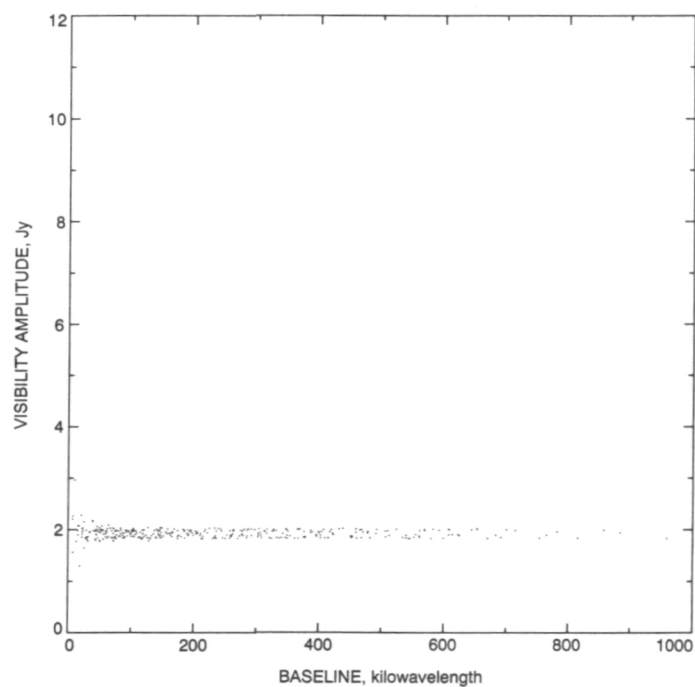


Fig. 8. Simulated visibility plot for observations of Galileo and Jupiter at their rise time (8 deg elevation) at the VLA on December 7, 1995. An 8-MHz bandwidth and use of the VLA A configuration are assumed.

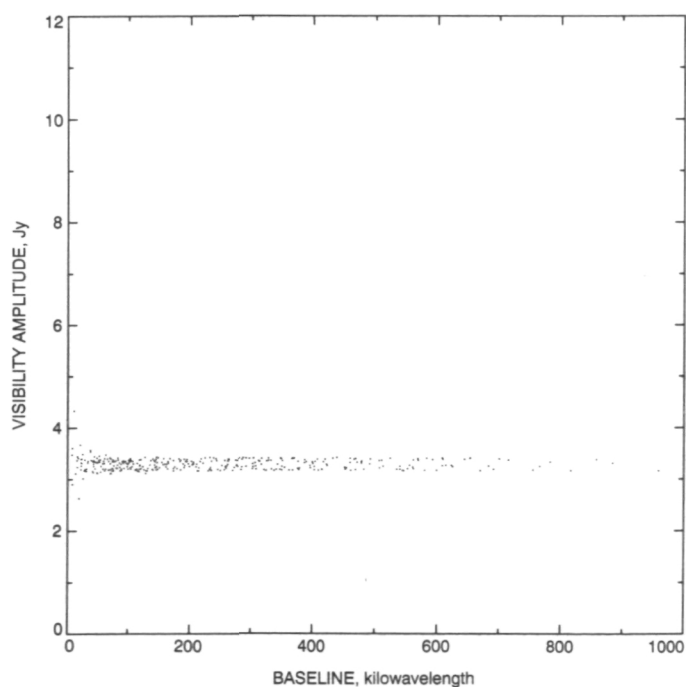


Fig. 10. Simulated visibility plot for observations of Galileo and Jupiter at their rise time (8 deg elevation) at the VLA on December 7, 1995. A 4.7-MHz bandwidth and use of the VLA A configuration are assumed.

1994025119

N94- 29622

TDA Progress Report 42-105

442612

May 15, 1991

Elevation Control System Model for the DSS 13 Antenna

W. Gawronski and J.A. Mellstrom
Ground Antennas and Facilities Engineering Section

In order to meet the requirements for precision pointing of 34-m antennas, adequate control design and simulation software have to be developed, with a detailed description of the supporting analytical tools. This article describes a control system model for the elevation drive of the DSS 13 antenna. The model allows one to simulate elevation dynamics, cross-coupled dynamics in azimuth and elevation, and RF pointing error. A modal state-space model of the antenna structure was obtained from its finite-element model with a free rotating tipping structure. Model reduction techniques were applied separately for the antenna model and rate-loop model, thereby reducing the system order to one-third of the original one while preserving its dynamic properties. Extensive simulation results illustrate properties of the model.

I. Introduction

In order to meet the requirements for the 34-m antenna pointing accuracy, an appropriate control system model has to be developed. The purpose of this article is to give a detailed description of the antenna control system modeling. This analytical approach is the background for the computer software development, which is used for the simulations of antenna dynamics due to control input, disturbances, and parameter variations, and is a tool for the control system design.

A state-space control system model for elevation pointing of the DSS 13 antenna is described. Special attention is paid to antenna structure modeling. A finite-element model of the antenna structure with a free rotating tipping structure is used to obtain its state-space model in modal coordinates. The model was developed with the following assumptions:

- (1) Linearity of the model (nonlinear effects due to dry friction are omitted).

- (2) The structural model consists of mode shapes up to 5 Hz.
- (3) The elevation angle of the tipping structure is 90 deg.
- (4) The azimuth rotation joint is locked.
- (5) No disturbances act on the structure.

The balanced reduction technique was applied to reduce the structural model and subsequently the rate-loop model. The order of the resulting reduced model is one-third of the original one, while the modeling error is insignificantly small. Extensive simulations were performed and are described below to illustrate the control system properties.¹

¹ W. Gawronski, "Software Package for Modeling and Simulation of the Elevation Control System for the DSS 13 Antenna," JPL Interoffice Memorandum 3324-91-016 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 6, 1991.

II. Structural Model

Here and throughout the article, the state-space representation of a system is used. This representation is denoted by a triple (A, B, C) , which represents a first-order differential equation in the form

$$\dot{x} = Ax + Bu, \quad y = Cx \quad (1)$$

Assuming the state vector x of dimension $n \times 1$, the input u of dimension $p \times 1$, and the output y of dimension $q \times 1$, the matrices A , B , and C are of dimensions $n \times n$, $n \times p$, and $q \times n$, respectively.

The state-space model for an antenna structure with a 90-deg elevation angle and locked azimuth rotation is obtained. A balanced reduction technique applied to the structural model gives the smallest acceptable model.

A. Full Model

In this section the state-space triple (A_s, B_s, C_s) for the antenna structure is determined. The triple is obtained in modal coordinates. The state vector of the structure x_s consists of modal displacements q_s and modal velocities v_s :

$$x_s^T = [q_s^T, v_s^T] \quad (2)$$

The antenna structural model is generated from the natural frequencies ω_i and mode shapes ϕ_i , obtained from the finite-element model generated by JPL/IDEAS, and presented in Appendix A. Modal damping of the structure ζ_i is assumed to be 0.5 percent, i.e., $\zeta_i = 0.005$ for $i = 1, \dots, 10$. The antenna is a free rotating structure with a joint between the alidade and tipping structure. Hence, the open-loop structural model is unstable. From the finite-element model, $m = 10$ modes are obtained; the first one is the rigid-body mode with zero natural frequency. The modes are determined for selected p points of interest. Thus, $\phi_i = [\phi_{i1}, \phi_{i2}, \dots, \phi_{ip}]^T$ for $i = 1, \dots, 10$.

1. Matrix A_s . Denote

$$\Omega = \text{diag}(\omega_i), \quad Z = \text{diag}(\zeta_i), \quad i = 1, 2, \dots, m \quad (3)$$

Then, the system matrix A_s for the antenna structure is

$$A_s = \begin{bmatrix} 0 & I_m \\ -\Omega^2 & -2Z\Omega \end{bmatrix} \quad (4)$$

where I_m is the identity matrix of dimensions $m \times m$. Details of the derivation of A_s are presented in Appendix B. The numerical values of Ω , Z , and other parameters are given in Appendix A.

2. Matrix B_s . The bull-gear-pinion connection is shown in Fig. 1. The rollers of the connection are redundant; they compensate for gravity forces. The rigid pinion housing includes two drive systems. A torque applied at the pinion of the gearbox is the input to the structure. The pinion radius is $r_p = 6$ inches. A node labeled $no = 86302$ in the finite-element model is a point of contact between the pinion and the bull gear, while a node with the label $nu = 86881$ is located at the joint of the supporting truss structure and the pinion housing. The joint nu is at distance ρ from node no . The distance ρ is small in comparison to the bull gear radius; thus, in further analysis the nodes nu and no coincide. The torque T applied to the pinion is in equilibrium with tangential (to the bull gear and the pinion) forces F_{no} and F_{nu} , see Fig. 2. For the force F_{no} with y and z components F_{noy} and F_{noz} , and the absolute value $|F_{no}|$, one obtains from Fig. 2

$$|F_{no}| = |T|/r_p,$$

$$F_{noy} = |F_{no}| \sin \alpha, \quad F_{noz} = |F_{no}| \cos \alpha$$

Hence,

$$F_{no} = \begin{bmatrix} F_{noy} \\ F_{noz} \end{bmatrix} = \begin{bmatrix} r_p^{-1} \sin \alpha \\ r_p^{-1} \cos \alpha \end{bmatrix} T = B_{no} T$$

where $B_{no} = r_p^{-1}[\sin \alpha \cos \alpha]^T$ and $\alpha = 28$ deg. From Fig. 2, one can see that $F_{nu} = -F_{no}$; hence, $F_{nu} = -B_{no}T$. In global coordinates, the force F_s acting on the structure is a vector $F_s^T = [0 \ 0 \ \dots \ F_{no}^T \ 0 \ \dots \ F_{nu}^T \ 0 \ \dots \ 0]$. Thus,

$$F_s = B_o T$$

where $B_o^T = [0 \ 0 \ \dots \ B_{no}^T \ 0 \ \dots \ -B_{no}^T \ 0 \ \dots \ 0]$. The matrix B_o is transformed into modal coordinates:

$$B_s = M_m^{-1} \Phi^T B_o \quad (5a)$$

where $M_m = \text{diag}(m_{m1}, \dots, m_{mm})$ is the modal mass matrix. Denoting $\phi_{no} = [\phi_{1no}, \dots, \phi_{mno}]$, $\phi_{nu} = [\phi_{1nu}, \dots,$

ϕ_{mnu}], $\phi_{ino}^T = [\phi_{inoy} \ \phi_{inoz}]$, and $\phi_{inu} = [\phi_{inuy} \ \phi_{inuz}]$, where ϕ_{inoy} is the i th mode component at node no , at y direction (mode shapes are listed in Appendix A), the above equation reduces to

$$B_s = M_m^{-1}(\Phi_{no} - \Phi_{nu})B_{no} \quad (5b)$$

3. Matrix C_s . The following outputs of the antenna structure are considered: elevation angle θ_e , elevation rate $\dot{\theta}_e$, pinion rate $\dot{\theta}$, pointing error angle in the x direction ε_x , and pointing error angle in the y direction ε_y . Hence, the antenna output matrix C_s consists of five rows: C_{s1} for the elevation angle, C_{s2} for the elevation rate, C_{s3} for the pinion rate, C_{s4} for pointing error in the x direction, and C_{s5} for pointing error in the y direction:

$$C_s^T = [C_{s1}^T \ C_{s2}^T \ C_{s3}^T \ C_{s4}^T \ C_{s5}^T] \quad (6)$$

The first two rows of C_s are determined as follows. In the finite-element model, the node at the bull gear center has a label $nb = 5380$, and the node at distance R to nb has a label $nc = 41212$ (see Fig. 2). High stiffness of the bull gear and the close location of the two nodes allows one to determine the elevation angle as a rigid-body rotation:

$$\theta_e = (y_{nc} - y_{nb})/R$$

$$= (C_{ncy}\Phi - C_{nby}\Phi)x_s/R = (\phi_{ncy} - \phi_{nby})x_s/R$$

where C_{ncy} and C_{nby} denote rows with all but one zero elements. The nonzero element of C_{ncy} is equal to 1, and is located in the position corresponding to the location of y displacement at node nc in the vector x_s . The nonzero element of C_{nby} is equal to 1, and located in the position corresponding to the location of y displacement at node nb in the vector x_s . Furthermore, $\phi_{ncy} = [\phi_{1ncy}, \dots, \phi_{mncy}]$, where ϕ_{incy} is the i th mode component at node nc at y direction, and $\phi_{nby} = [\phi_{1nby}, \dots, \phi_{mnby}]$, where ϕ_{inby} is the i th mode component at node nb at y direction. From the above, it follows that

$$C_{s1} = [0 \ C_{se}], \quad C_{s2} = [C_{se} \ 0] \quad (7a)$$

where $C_{se} = (\phi_{ncy} - \phi_{nby})/R$.

For the determination of the pinion rate measurement matrix C_{s3} , denote the velocity at pinion housing v_p (at node nu), bull gear velocity v_b at node no (see Fig. 2),

and their y and z components by v_{py} , v_{pz} , v_{by} , and v_{bz} , respectively. Projections of v_b and v_p onto directions tangential to the bull gear at node no are denoted v_1 and v_2 , respectively.

$$v_1 = v_{by} \sin \alpha + v_{bz} \cos \alpha, \quad v_2 = v_{py} \sin \alpha + v_{pz} \cos \alpha$$

The pinion rate is $\dot{\theta} = (v_1 - v_2)/r_p$, and noting that $v_{by} = \phi_{noy}v_s$, $v_{bz} = \phi_{noz}v_s$, $v_{py} = \phi_{nuy}v_s$, and $v_{pz} = \phi_{nuz}v_s$, where v_s is a modal velocity defined in Eq. (2), one obtains

$$v_1 - v_2 = [(\phi_{noz} - \phi_{nuz}) \cos \alpha + (\phi_{noy} - \phi_{nuy}) \sin \alpha]v_s$$

Therefore, $\dot{\theta} = C_{sp}v_s = [0 \ C_{sp}]x_s$, giving finally

$$C_{s3} = [0 \ C_{sp}] \quad (7b)$$

where $C_{sp} = [(\phi_{noz} - \phi_{nuz}) \cos \alpha + (\phi_{noy} - \phi_{nuy}) \sin \alpha]/r_p$.

The last two rows of C_s , the output matrices for the pointing error in the x and y directions, are:

$$C_{s4} = [C_{ex} \ 0], \quad C_{s5} = [C_{ey} \ 0] \quad (7c)$$

where $C_{ex} = [333.26, 0.07973, -0.60701, 0.15431, 48.635, 4.0901, -40.343, 113.24, 225.17, -39.59]$ and $C_{ey} = [-9.0557, -24.706, -6.6233, 3.6844, -19.179, 23.571, 43.871, 180.87, -59.909, -227.56]$ are obtained from the JPL/IDEAS.

B. Model Reduction

The structural model under consideration consists of 10 modes or 20 states. The set of modes may consist of modes which are not important in system dynamics and can thus be eliminated. Observability and controllability properties are used to determine which modes to eliminate. Observability is a property of a system that indicates the participation of each system state in the output with all states equally excited. Controllability, on the other hand, indicates the excitation of each state due to white noise, or impulse input. A balanced representation [1] is a state-space representation with its states equally controllable and observable, and the Hankel singular value is the measure of the joint controllability and observability of each balanced state variable. The states with small Hankel singular values can be deleted, since they are weakly excited and weakly observed at the same time, causing minimal

modeling error. For flexible structures, modal representation is almost balanced, assuming small damping and distinct poles [2,3]. Each mode has almost the same controllability and observability property, hence, each mode can be considered for reduction separately. The Hankel singular value for the i th mode is given by Eq. (53) of [2] and Eq. (14) of [4]:

$$\gamma_i = \frac{\overline{(b_{si}b_{si}^T)(c_{si}^T c_{si})}}{4\zeta_i\omega_i^2} \quad (8)$$

where b_{si} is the i th row of B_s and c_{si} is the i th column of C_s . For the rigid-body mode, the Hankel singular value does not exist—its value tends to infinity—hence, the rigid-body mode is always included in the reduced model. For each mode, the Hankel singular value is determined to help the designer decide on the number of modes in the reduced structural model.

Hankel singular values of the 10-mode model of the antenna (A_s, B_s, C_s) are plotted in Fig. 3, where the Hankel singular value of the rigid-body mode is omitted. The reduced-order model consists of five modes: the rigid-body mode and four elastic modes with natural frequencies 3.1358, 4.1149, 4.4503, and 4.5688 Hz, indicated in Fig. 3. The transfer functions of the full and reduced models of the antenna are shown in Fig. 4, indicating that the reduced model fully represents properties of the original system. Note that disturbances are not included in the model; therefore, they are not considered in model reduction. One should note, however, that a different reduced model would be obtained for a model with disturbances included.

III. Rate-Loop Model

This section presents details of rate-loop modeling and introduces model reduction.

A. Full Model

The rate-loop model of the antenna is shown in Fig. 5. It consists of two identical subsystems denoted G_o , antenna structure G_s , command amplifier G_c , and the bias loop, with the bias command v_{bias} .

The subsystem G_c is shown in Fig. 6(a). The first state variable is denoted $x_1 = v_1$ and its state equation is

$$\dot{x}_1 = -\tau_1^{-1}x_1 + k_1k_s\tau_1^{-1}\dot{\theta}_o \quad (9)$$

The subsystem G_o , Fig. 6(b), consists of four subsystems itself: G_1, G_2 (both amplifiers), G_3 (motor armature), and G_4 (motor and gearbox mechanical properties). The state equations for G_1 are

$$\dot{x}_2 = x_3 \quad (10a)$$

$$\dot{x}_3 = x_1 - \tau_3^{-1}x_3 - k_{tach}x_7 \quad (10b)$$

where $x_7 = \omega_m$. Similarly, for G_2 :

$$\dot{x}_4 = x_5 \quad (11a)$$

$$\dot{x}_5 = k_r\tau_3^{-1}x_2 + k_r\tau_2\tau_3^{-1}x_3 - \tau_5^{-1}x_5 - k_{cur}x_6 \quad (11b)$$

where $x_6 = i_o$. For G_3 , from Fig. 6:

$$\begin{aligned} \dot{x}_6 &= k_f k_i L_a^{-1} \tau_5^{-1} x_4 + k_f k_i \tau_4 L_a^{-1} \tau_5^{-1} x_5 \\ &\quad - R_a L_a^{-1} x_6 - k_b L_a^{-1} x_7 \end{aligned} \quad (12a)$$

$$T_o = k_m x_6 \quad (12b)$$

The system G_4 denotes the motor and gearbox model. There are two inputs to the system: motor torque T_o and pinion rate $\dot{\theta}$. The output is the pinion torque T . The gearbox flexibility is described either by the input or the output stiffness, as shown in Figs. 7(a) and 7(b). For the output stiffness model, with notations as in Fig. 7(a), one obtains equations for the gearbox rate

$$T_N = NT_o, \quad \theta_N = \theta_m/N$$

and the dynamic equilibrium equations

$$J_m \dot{\omega}_m + T_N/N = T_o$$

along with the spring equation for the output stiffness k_{go}

$$T = T_N = k_{go}(\theta_m - N\theta)/N$$

Similar equations are derived for the input stiffness model in Fig. 7(b). Both models are represented by Fig. 8, where $N_x = 1/N$ and $k_{gx} = k_{go}$ for the output stiffness

model and $N_x = N$ and $k_{gx} = k_{gi}$ for the input stiffness model. Denoting $x_8 = T$, the state equations are

$$\dot{x}_7 = k_m J_m^{-1} x_6 - J_m^{-1} N^{-1} x_8 \quad (13a)$$

$$\dot{x}_8 = k_{gx} N_x x_7 - k_{gx} N_x N C_s x_s \quad (13b)$$

where $N_x = N$ for $k_{gx} = k_{gi}$, and $N_x = 1/N$ for $k_{gx} = k_{go}$.

Combining Eqs. (10)–(13) for models G_1 , G_2 , G_3 , and G_4 , one obtains the state-space model for G_o (in-

puts v_1 and $\dot{\theta}$, outputs T and i_o , and state vector $x_o = [x_2, x_3, x_4, x_5, x_6, x_7, x_8]^T$):

$$\dot{x}_o = A_1 x_o + B_{11} v_1 + B_{12} \dot{\theta} \quad (14a)$$

$$\begin{bmatrix} T \\ i_o \end{bmatrix} = C_1 x_o \quad (14b)$$

where

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1/\tau_3 & 0 & 0 & 0 & -k_{tach} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ k_r/\tau_3 & k_r \tau_2/\tau_3 & 0 & -1/\tau_5 & -k_{cur} & 0 & 0 \\ 0 & 0 & k_f k_i/L_a \tau_5 & k_f k_i \tau_4/L_a \tau_5 & -R_a/L_a & -k_b/L_a & 0 \\ 0 & 0 & 0 & 0 & k_m/J_m & 0 & -1/J_m N \\ 0 & 0 & 0 & 0 & 0 & k_{gx} N_x & 0 \end{bmatrix}$$

$$B_{11}^T = [0 \ 1 \ 0 \ 0 \ 0 \ 0 \ 0], \quad B_{12}^T = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ -k_{gx} N_x N], \quad C_{11} = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1],$$

$$C_{12} = [0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0], \quad C_1 = \begin{bmatrix} C_{11} \\ C_{12} \end{bmatrix}, \quad B_1 = [B_{11} \ B_{12}]$$

Two systems' G_o 's are located in the rate-loop system (see Fig. 5): the first with inputs v_{11} and $\dot{\theta}$, outputs T_1 and i_{o1} , and state vector x_{21} , and the second with inputs v_{12} and $\dot{\theta}$, outputs T_2 and i_{o2} , and state vector x_{22} . The equations for these systems are

$$\dot{x}_{21} = A_1 x_{21} + B_{11} v_{11} + B_{12} \dot{\theta} \quad (15a)$$

$$\begin{bmatrix} T_1 \\ i_{o1} \end{bmatrix} = C_1 x_{21} \quad (15b)$$

and

$$\dot{x}_{22} = A_1 x_{22} + B_{11} v_{12} + B_{12} \dot{\theta} \quad (16a)$$

$$\begin{bmatrix} T_2 \\ i_{o2} \end{bmatrix} = C_1 x_{22} \quad (16b)$$

The subsystem G_c is an amplifier for the bias signal, with input Δv_c and output v_u . Denoting $x_a = v_u$, its state space equation is obtained:

$$\dot{x}_a = -\tau_6^{-1} x_a + k_{ctfr} \tau_6^{-1} \Delta v_c \quad (17)$$

The antenna structure equations according to Section II are

$$\dot{x}_s = A_s x_s + B_s T, \quad \dot{\theta}_e = C_{s2} x_s, \quad \dot{\theta} = C_{s3} x_s \quad (18)$$

The nodal equations are

$$T = T_1 + T_2, \quad T_1 = C_{11}x_{21}, \quad T_2 = C_{11}x_{22} \quad (19a)$$

$$\Delta v_c = v_{c1} - v_{c2} + k_{bs}v_{bias} = k_{ct}C_{12}(x_{21} - x_{22}) + k_{bs}v_{bias} \quad (19b)$$

$$v_{11} = v_1 - v_u = x_1 - x_a, \quad v_{12} = v_1 + v_u = x_1 + x_a \quad (19c)$$

Denoting the state vector $x_r^T = [x_1^T, x_{21}^T, x_{22}^T, x_a^T, x_s^T]$ and combining Eqs. (9)–(19), the equations for the whole system can be written

$$\dot{x}_r = A_r x_r + B_r u, \quad y = C_r x_r \quad (20)$$

where $u^T = [\dot{\theta}_o \ v_{bias}]$ and $y^T = [\theta_e^T \ \dot{\theta}_e^T \ \ddot{\theta}^T]$, and

$$A_r = \begin{bmatrix} -1/\tau_1 & 0 & 0 & 0 & 0 \\ B_{11} & A_1 & 0 & -B_{11} & B_{12}C_{s3} \\ B_{11} & 0 & A_1 & B_{11} & B_{12}C_{s3} \\ 0 & p_1 & -p_1 & -1/\tau_6 & 0 \\ 0 & B_s C_{11} & B_s C_{11} & 0 & A_s \end{bmatrix}$$

$$B_r^T = \begin{bmatrix} k_1 k_s / \tau_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p_2 & 0 \end{bmatrix},$$

$$C_r = \begin{bmatrix} 0 & 0 & 0 & 0 & C_{s1} \\ 0 & 0 & 0 & 0 & C_{s2} \\ 0 & 0 & 0 & 0 & C_{s3} \end{bmatrix},$$

$$p_1 = k_{ctfr} k_{ct} C_{12} / \tau_6, \quad p_2 = k_{ctfr} k_{ct} / \tau_6$$

B. Model Reduction

The rate-loop system model is reduced by applying the balancing principle. The system is transformed into balanced coordinates where each state is equally controllable and observable and not correlated with other states. The

balancing procedure from [3] is used, and is summarized in Appendix C. The procedure from Section II is applicable to flexible structures only.

The rate-loop system, with already reduced antenna structural modes, has 26 state variables. A plot of the system Hankel singular values is shown in Fig. 9. The Hankel singular values of component number 15 and larger are small in comparison to the remaining components. Therefore, the reduced system consists of 14 state variables. The transfer functions of the reduced and full rate-loop systems are compared in Fig. 10. The full system contains the full structural model: 36 state variables altogether. From Fig. 10 one can see that the reduced model of 14 state variables almost exactly approximates the full 36-state model. In Fig. 11 step responses of full and reduced models are compared, showing good agreement between the full and reduced models.

IV. Position-Loop Model

The rate-loop system, with the position loop closed, is shown in Fig. 12. A proportional-plus-integral (PI) controller is applied. For the state x_o defined $x_o^T = [x_r^T \ x_i]$, the state-space triple of the series connection of the rate-loop system and the controller is

$$A_o = \begin{bmatrix} A_r & B_t k_{ii} \\ 0 & 0 \end{bmatrix}, \quad B_o = \begin{bmatrix} B_t k_{pp} \\ 1 \end{bmatrix}, \quad (21)$$

$$C_o = [C_t \ 0], \quad B_{ob} = \begin{bmatrix} B_{bi} \\ 0 \end{bmatrix}$$

where k_{pp} and k_{ii} are the proportional and integral parameters of the controller, B_t is the first column of B_r , B_{bi} is the second column of B_r , and C_t is the first row of C_r . The closed-loop system matrix is

$$A_{cl} = A_o - B_o C_o \quad (22)$$

and the closed-loop triple is (A_{cl}, B_o, C_o) . The closed-loop system performance is illustrated in Fig. 13 in time domain and in Fig. 14 in frequency domain. One can see that the x component of the RF pointing error is much larger than the component in the y direction, and the y component of subreflector vertex displacement is much larger than the component in the x direction. In fact, for a symmetric

structure the RF y component and x displacement of the subreflector vertex should be zero.

Finally, the action of the bias input v_{bias} is tested. Two step inputs are introduced to the closed-loop system: a v_{bias} step of amplitude 2.5 V at the moment $t = 0$ sec and a command step θ_o of amplitude 0.1 deg at the moment $t = 10$ sec. The results are shown in Figs. 15(a) through 15(c). For the first 10 sec, when the command θ_o is zero, the torques T_1 and T_2 are equal but with opposite sign, so that the resulting torque T acting on the structure is zero: v_{bias} does not influence the structure. Nevertheless, both motors are active, their rates ω_{m1} and ω_{m2} are nonzero, plotted in Fig. 15(c). After 10 sec, when the command is active, the system behavior is similar to the one without bias.

V. Conclusions

In this article, modeling techniques for the DSS 13 antenna structure and its rate-loop control system have been

presented. Through balancing the system controllability and observability properties, a reduced system model has been obtained. Antenna dynamics, rate-loop dynamics, and dynamics of a system with a closed position loop have been simulated. The model allows one to simulate elevation dynamics, cross-coupled dynamics in azimuth and elevation, and a dynamic RF pointing error.

Section I contains restrictions that need to be relaxed. The structural model could include modes of up to 10 Hz, with free elevation and azimuth joints. The elevation and azimuth control systems could be simulated jointly, exposing the cross-coupling of the two systems. Disturbances (wind and measurement noise) could be included in the model. The system could be modeled for several elevation angles of the range of [0 deg, 90 deg], and the dependence of the system dynamics on elevation angle investigated. Nonlinear effects due to dry friction and saturations could be included. Finally, field measurements need to be performed, and the model parameters adjusted such that a satisfactory agreement between simulations and field data is achieved.

Acknowledgments

The authors thank Robert Hill, Roy Levy, Ben Parvin, and Douglas Strain for their extensive and useful discussions.

References

- [1] B. C. Moore, "Principal Component Analysis in Linear Systems: Controllability, Observability and Model Reduction," *IEEE Trans. Autom. Control*, vol. 26, no. 1, pp. 17-32, January 1981.
- [2] W. Gawronski and J.-N. Juang, "Model Reduction for Flexible Structures," in *Control and Dynamics Systems*, ed. C. T. Leondes, vol. 36, New York: Academic Press, 1990.
- [3] E. A. Jonckheere, "Principal Component Analysis of Flexible Systems—Open Loop Case," *IEEE Trans. Autom. Control*, vol. 29, no. 12, pp. 1095-1097, December 1984.
- [4] W. Gawronski and T. Williams, "Model Reduction for Flexible Space Structures," *Journal of Guidance, Control, and Dynamics*, vol. 14, no. 1, pp. 68-76, January 1991.

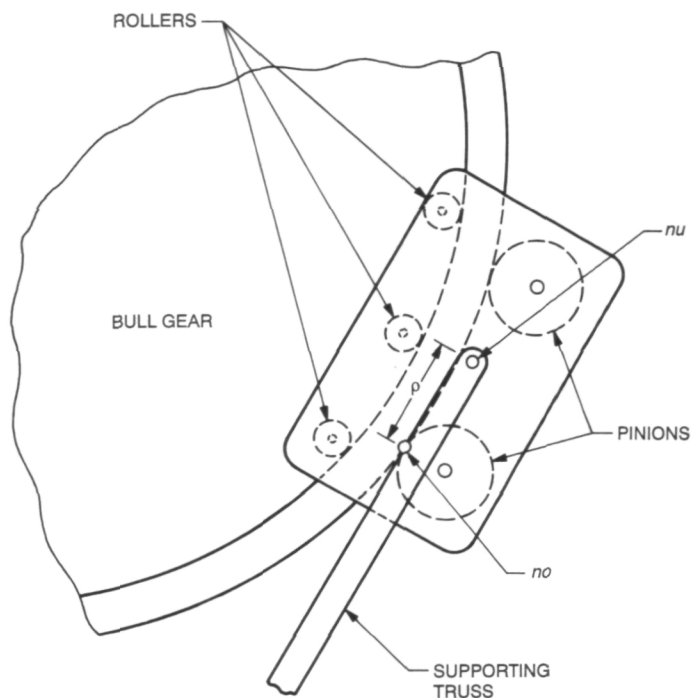


Fig. 1. Bull-gear-pinion connection.

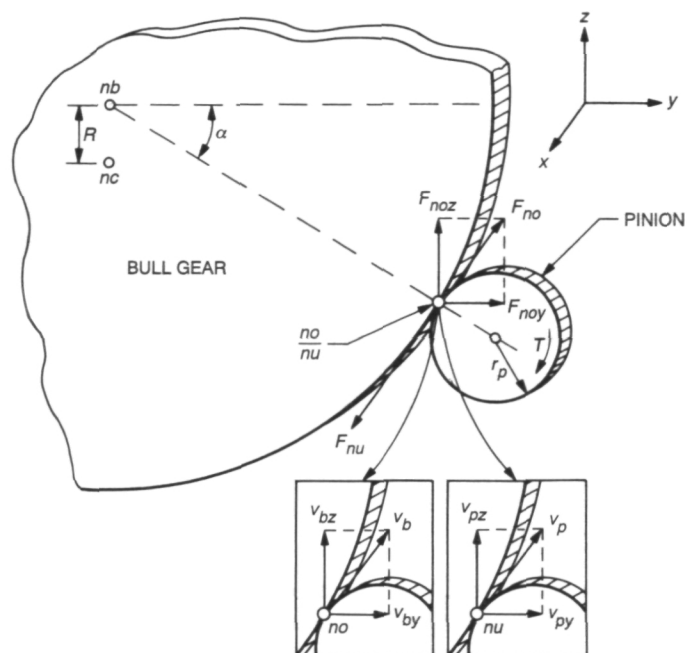


Fig. 2. Forces and velocities at the bull-gear-pinion connection.

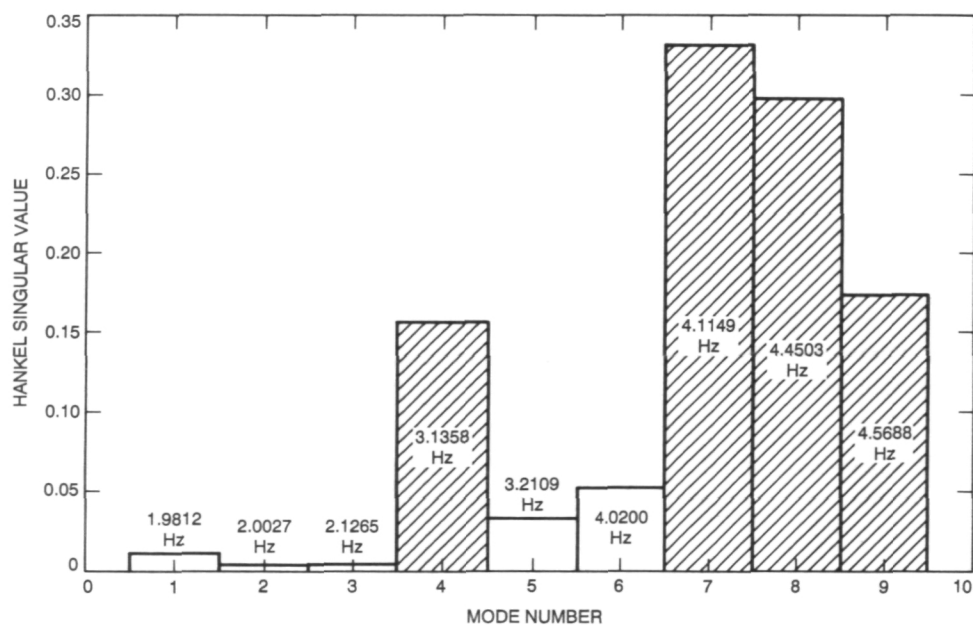


Fig. 3. Hankel singular values for the antenna structure.

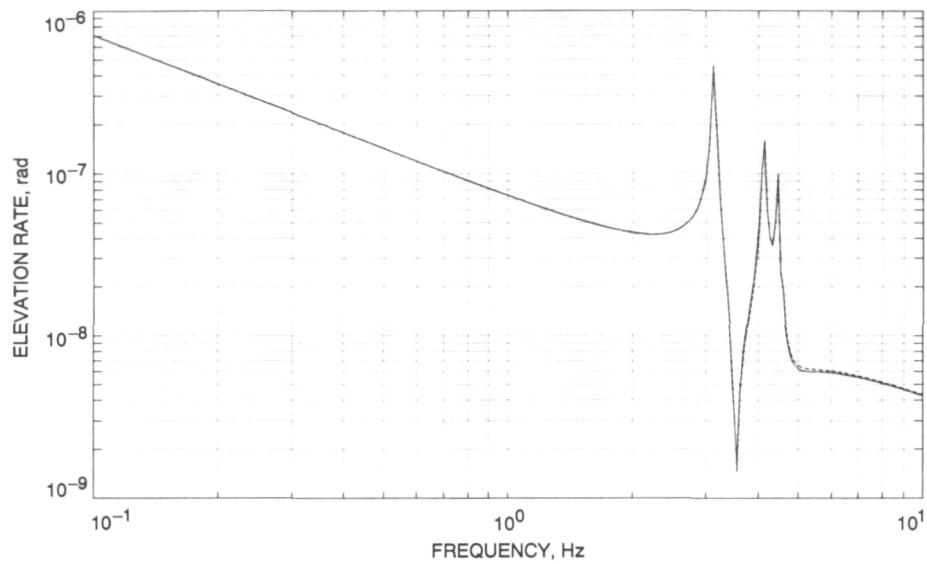


Fig. 4. Magnitude of the elevation rate transfer function for the full 10-mode (solid line) and the reduced 5-mode (dashed line) models of the antenna structure.

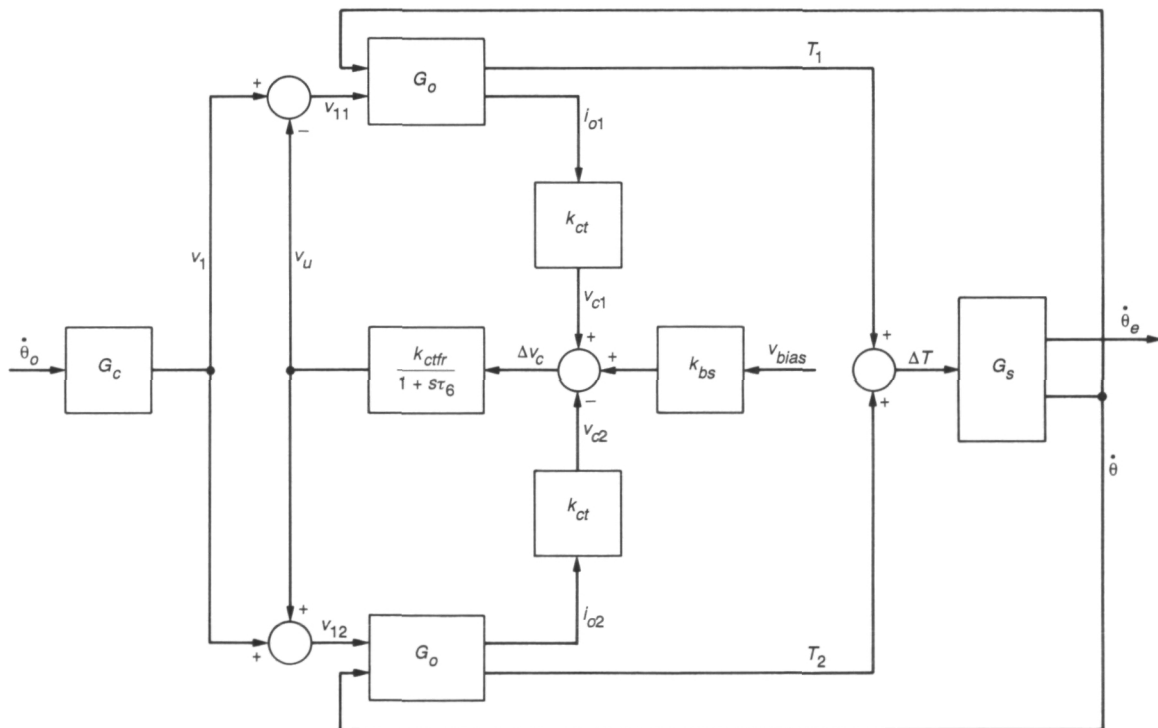


Fig. 5. Rate-loop control system.

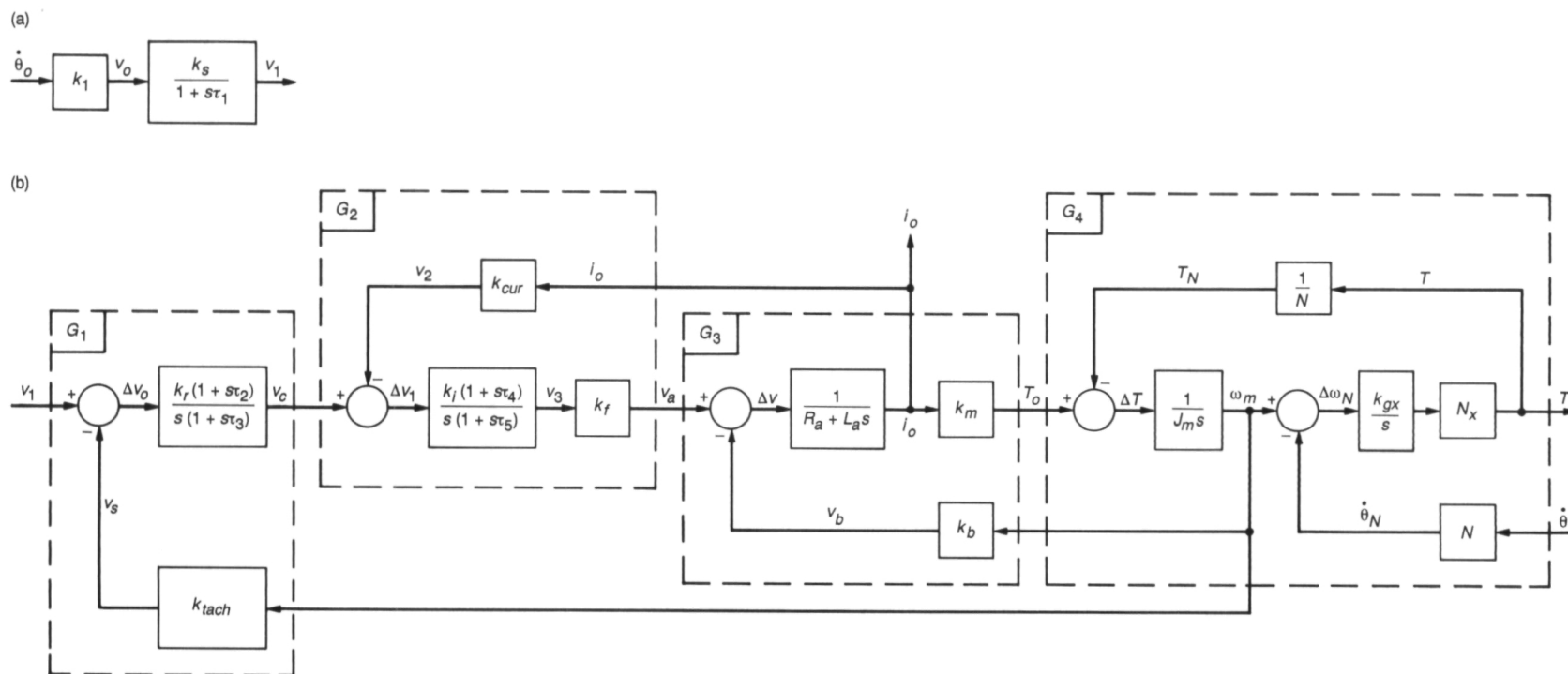


Fig. 6. Rate-loop subsystems: (a) subsystem G_c , and (b) subsystem G_o .

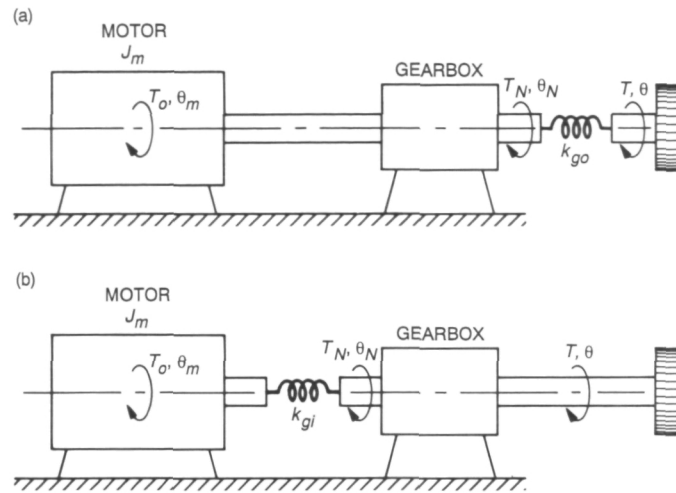


Fig. 7. Motor and gearbox configuration: (a) output stiffness model, and (b) input stiffness model.

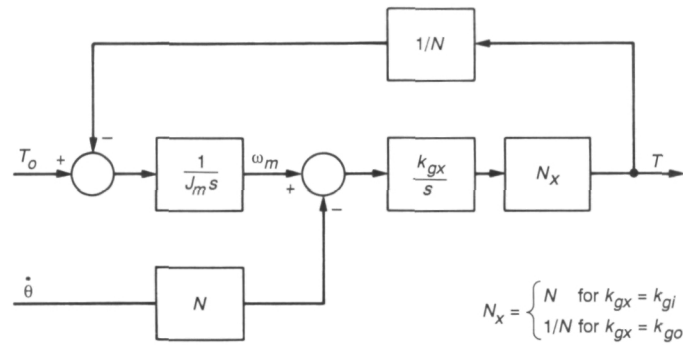


Fig. 8. Motor and gearbox model.

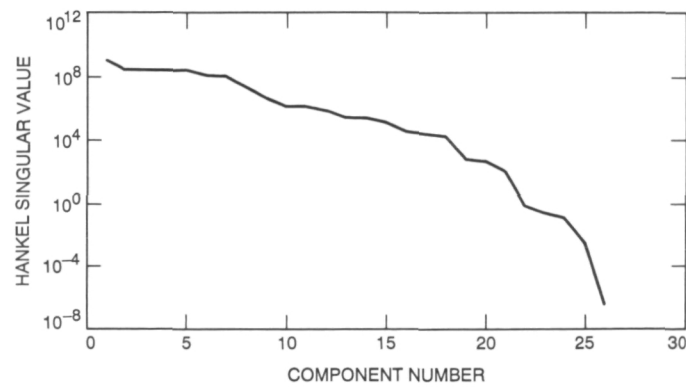


Fig. 9. Hankel singular values for the rate-loop system.

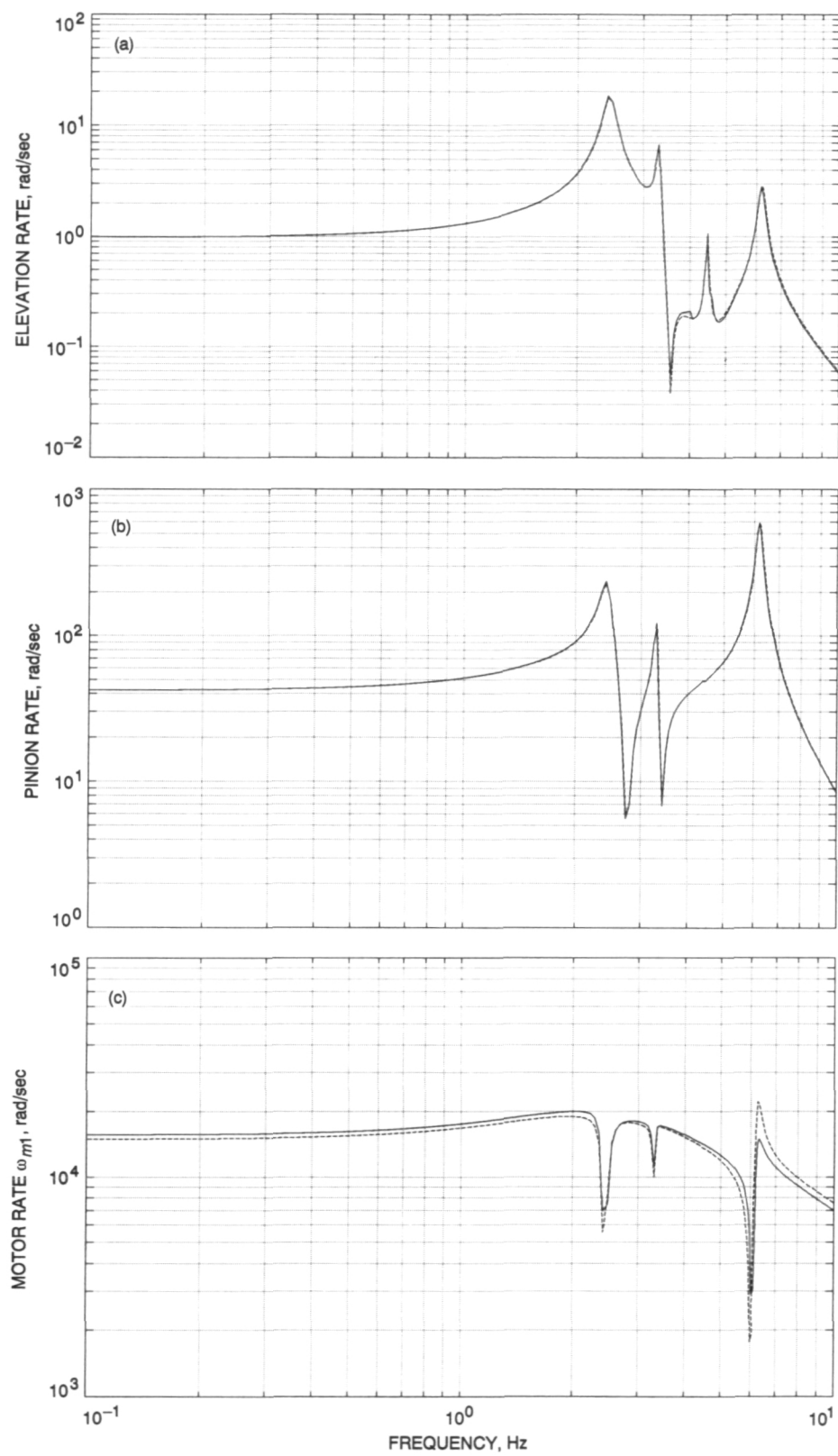


Fig. 10. Magnitudes of transfer function for the full 36-state (solid line) and the reduced 14-state (dashed line) rate-loop models (input: elevation rate, rad/sec).

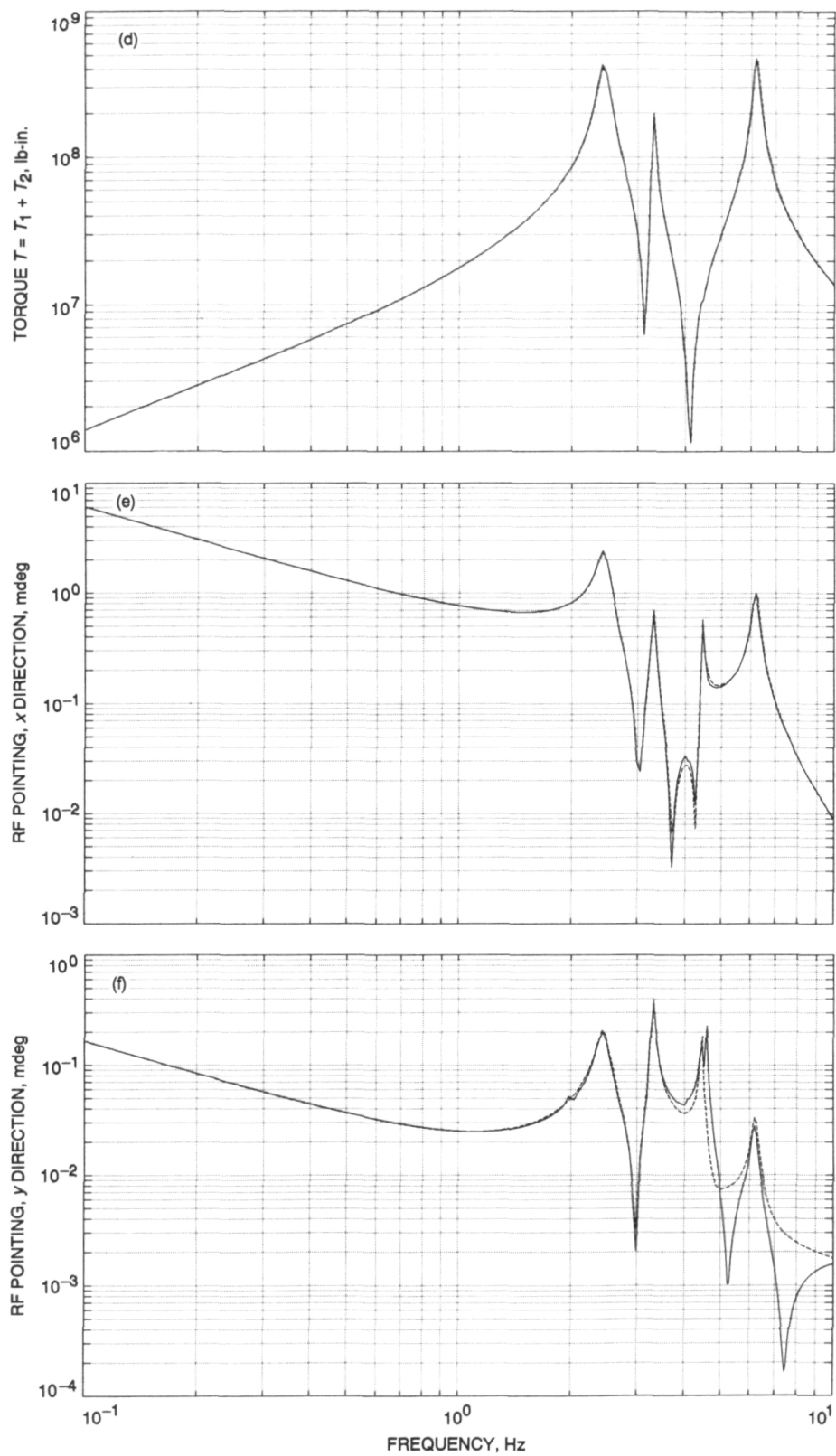


Fig. 10 (contd)

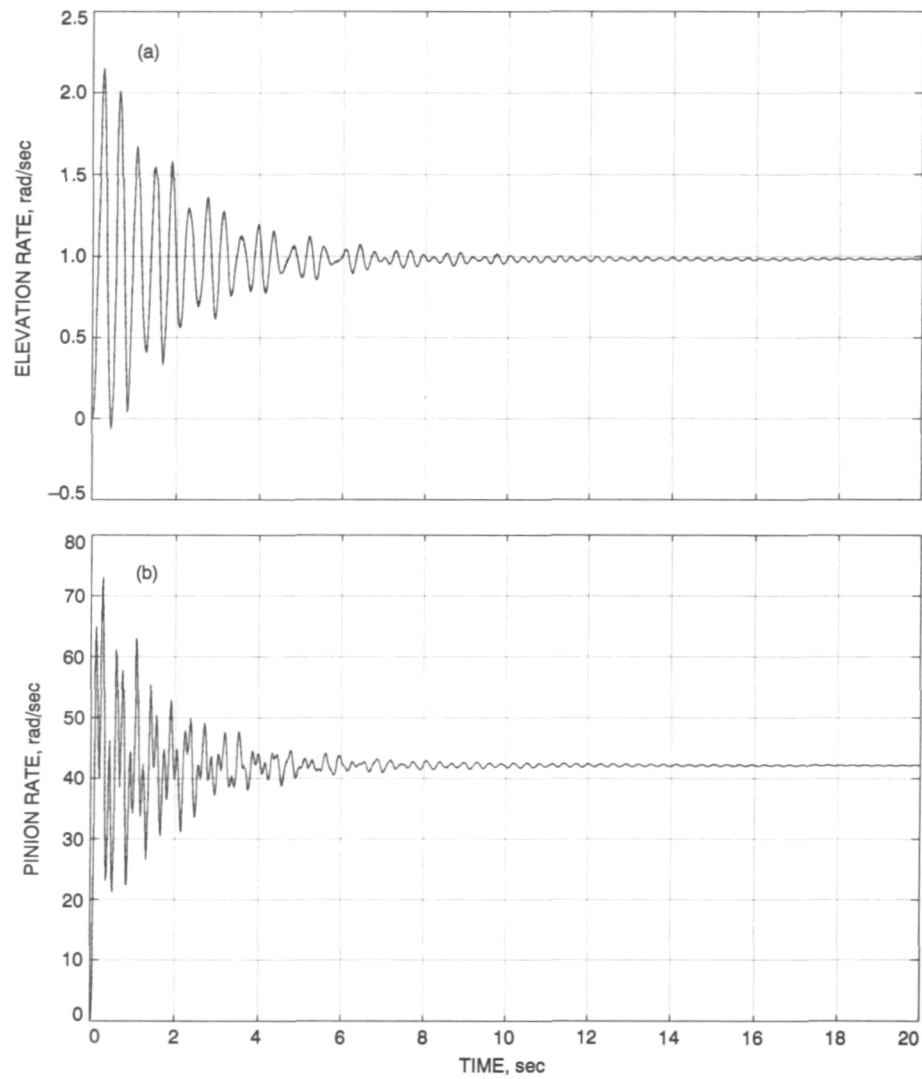


Fig. 11. Step responses of the full 36-state (solid line) and reduced 14-state (dashed line) rate-loop models (Input: elevation rate, rad/sec).

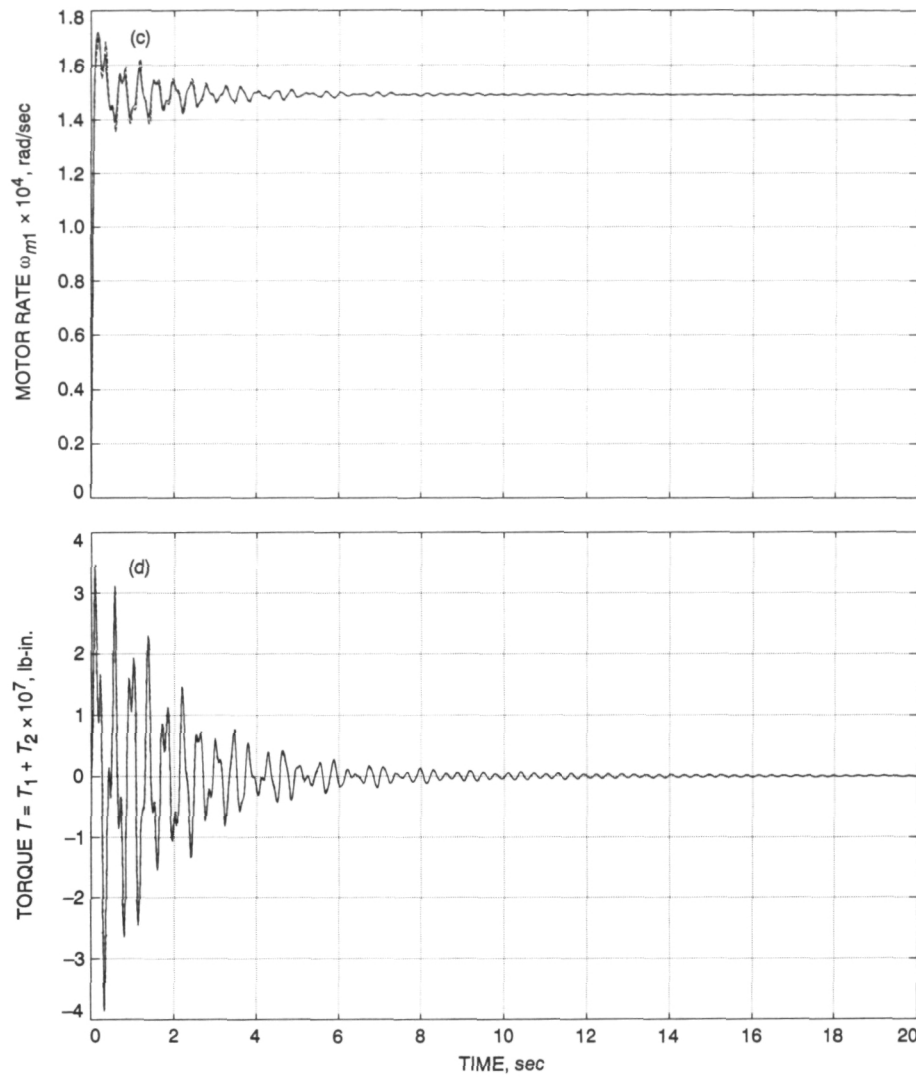


Fig. 11 (contd)

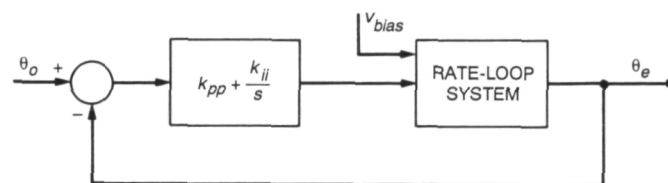


Fig. 12. Position-loop system.

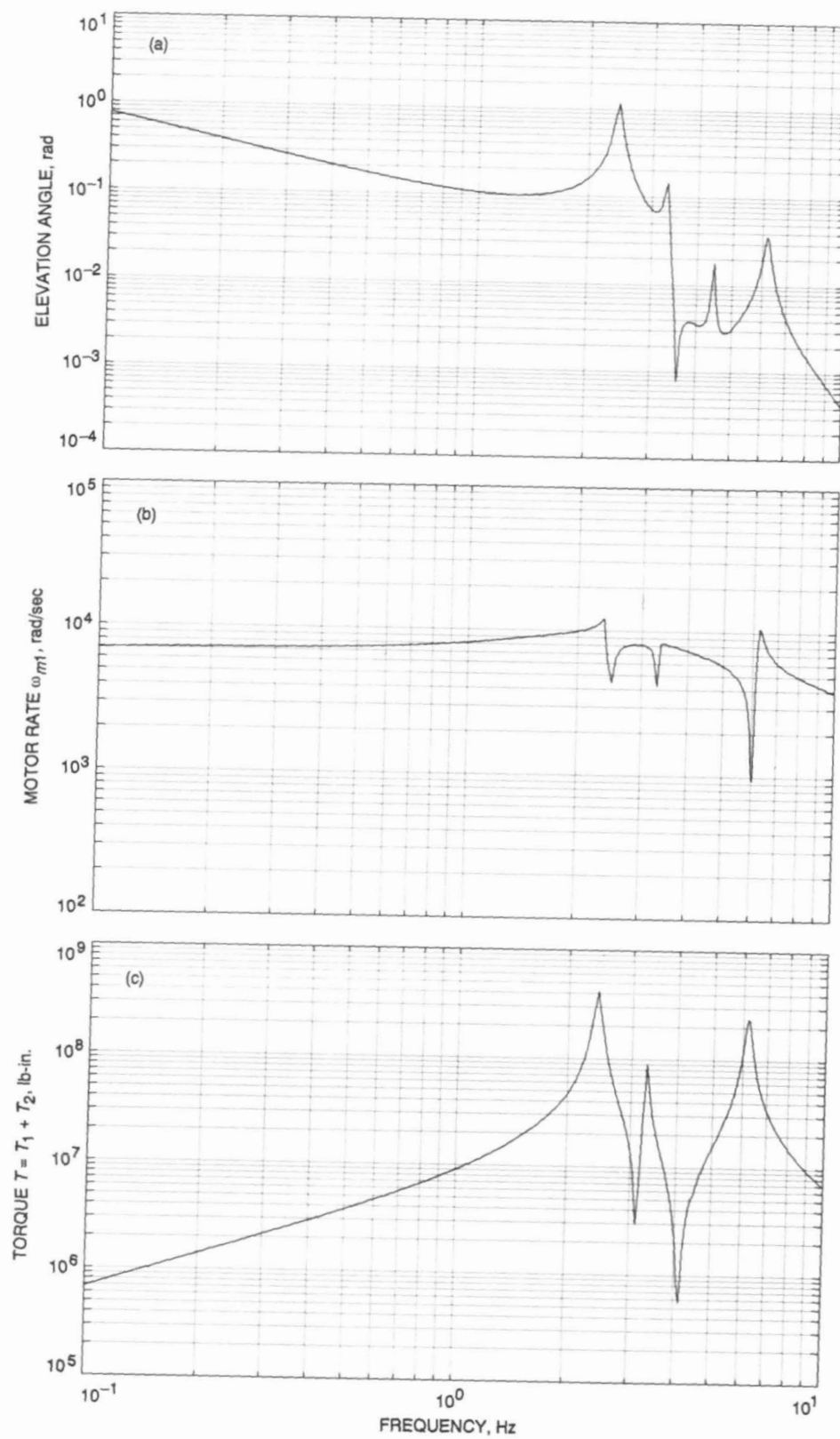


Fig. 13. Magnitudes of transfer function for the position-loop system; position loop closed, $k_{pp} = 0.5$, $k_{ii} = 0.1$; input: elevation angle, rad.

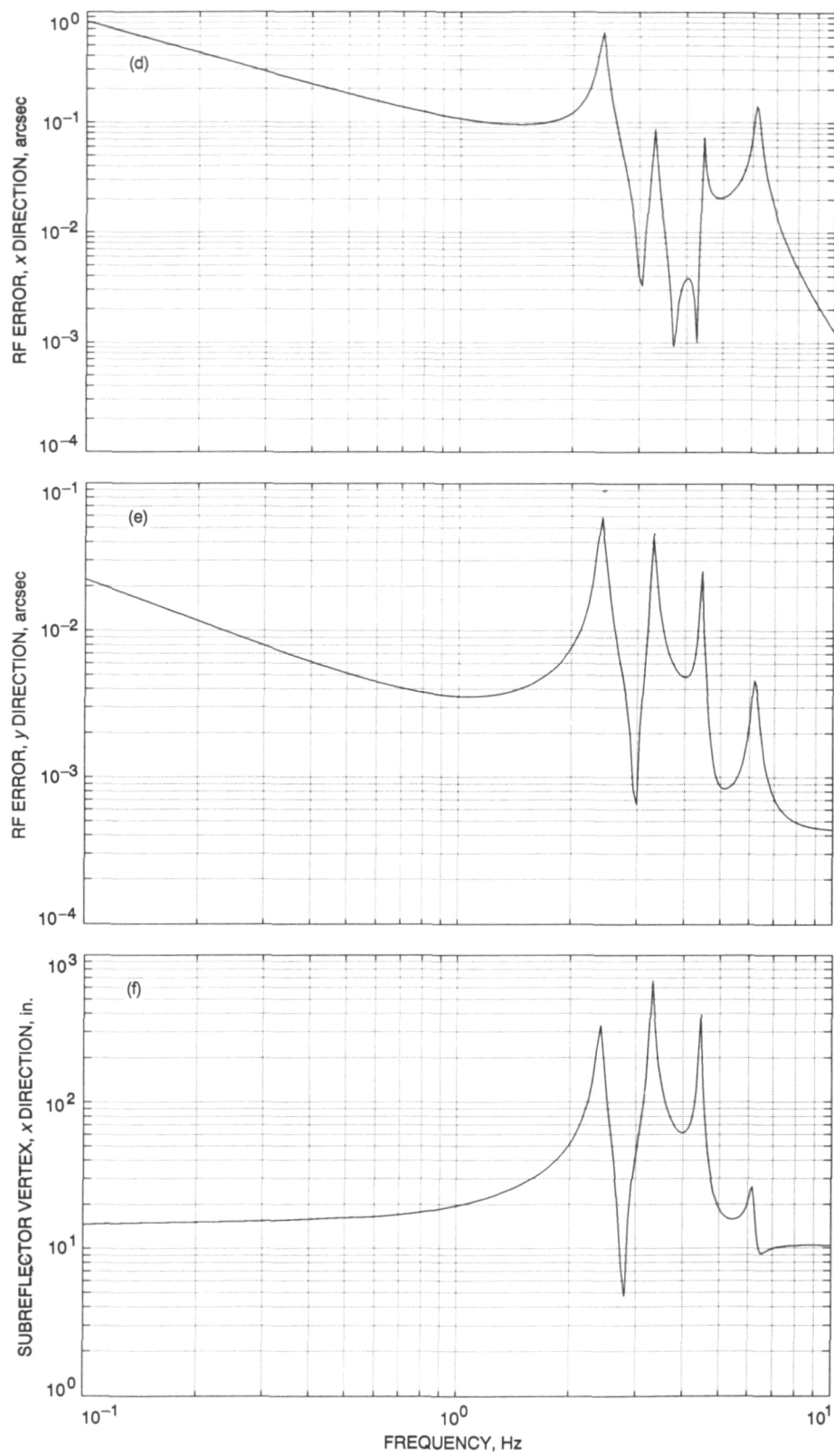


Fig. 13 (contd)

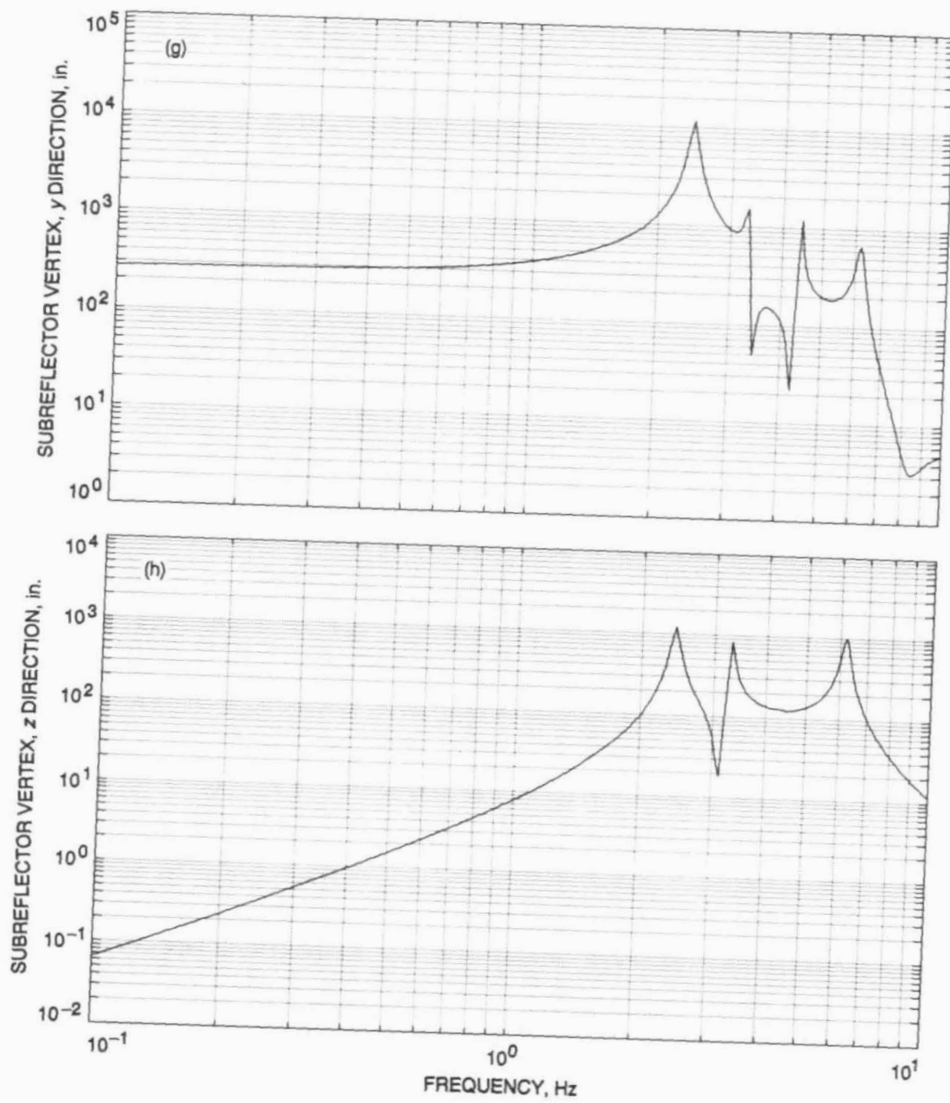


Fig. 13 (contd)

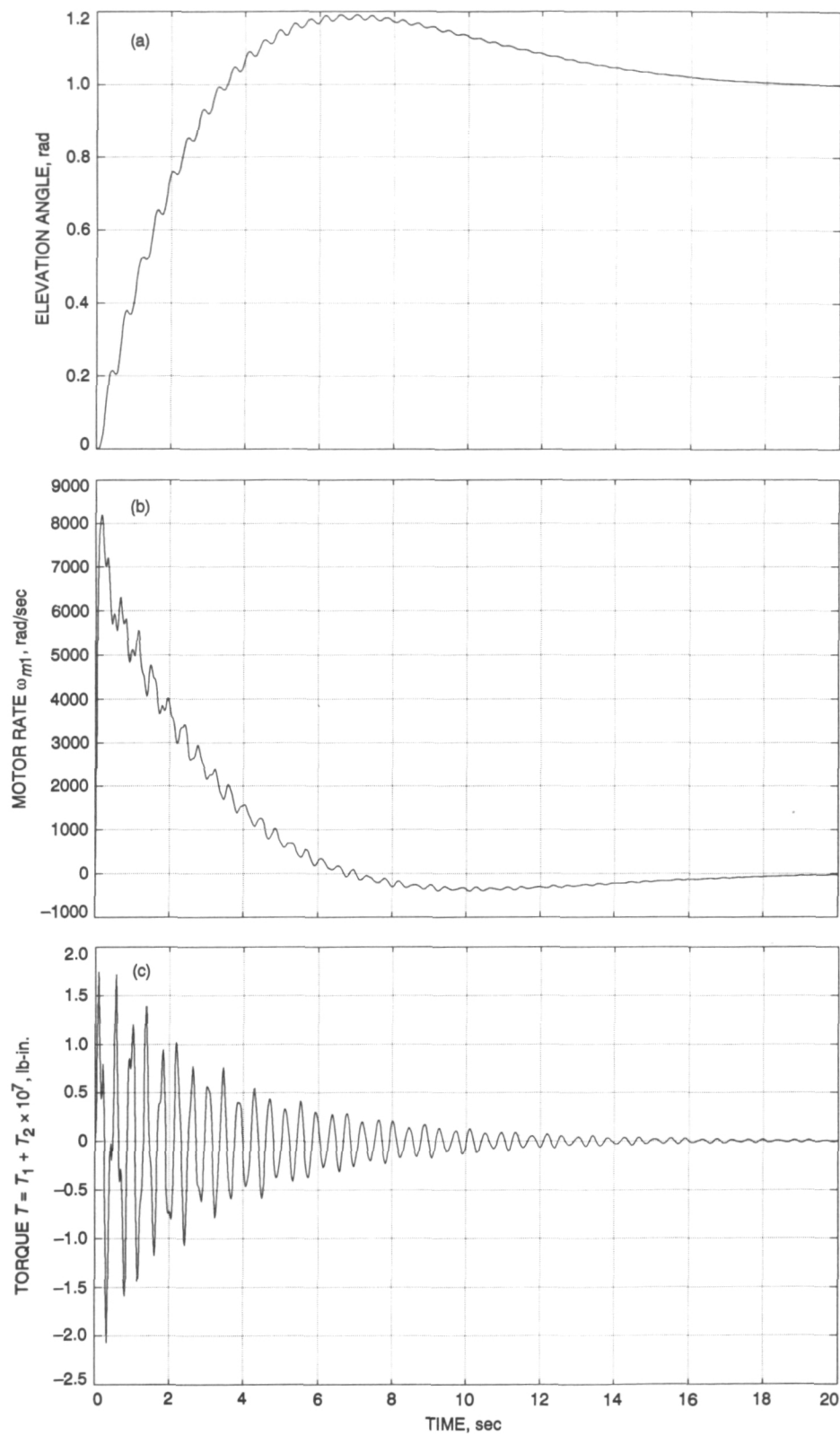


Fig. 14. Step responses for the closed-loop system; position loop closed, $k_{pp} = 0.5$, $k_{ij} = 0.1$; input: elevation angle, rad.

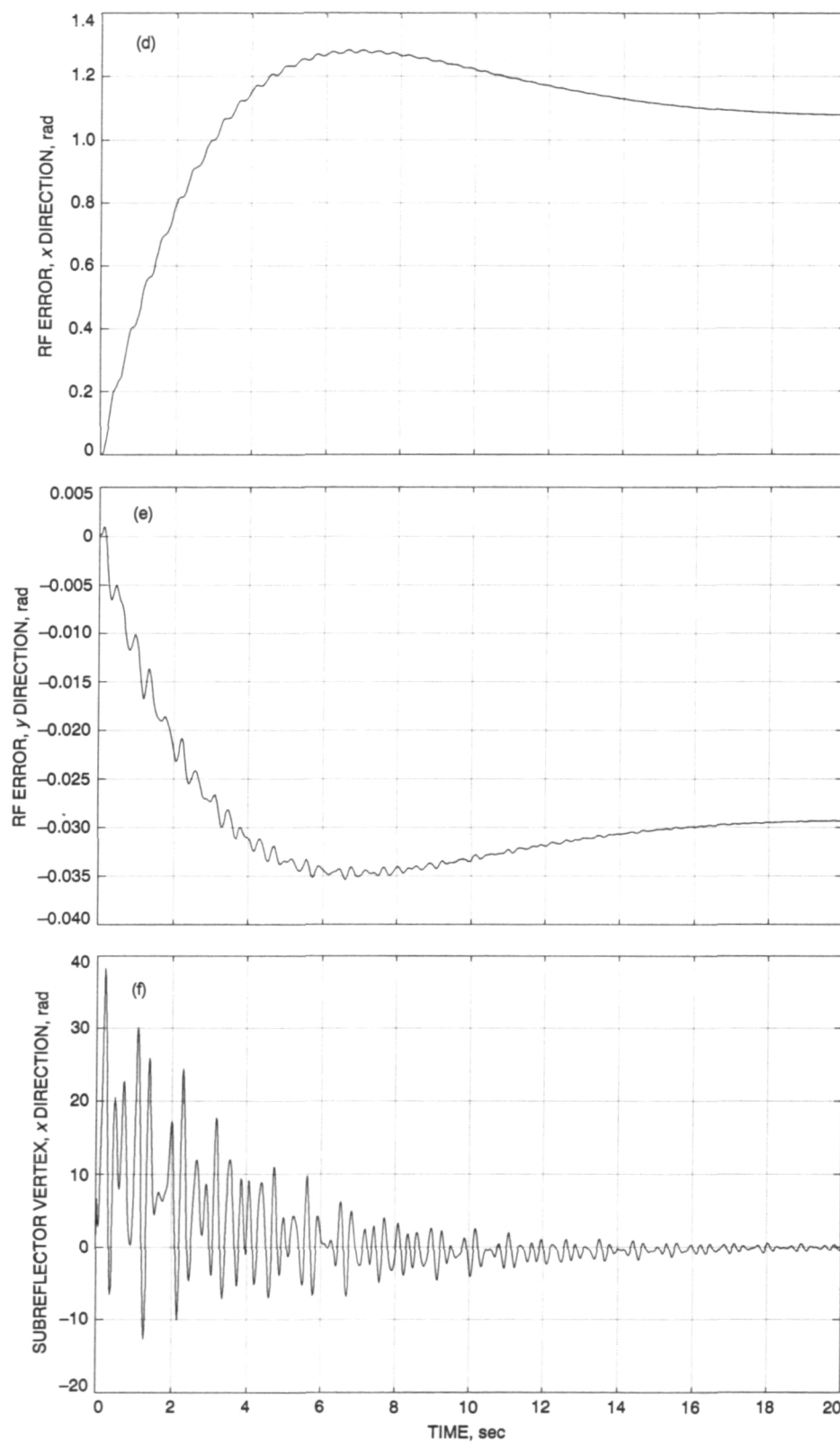


Fig. 14 (contd)

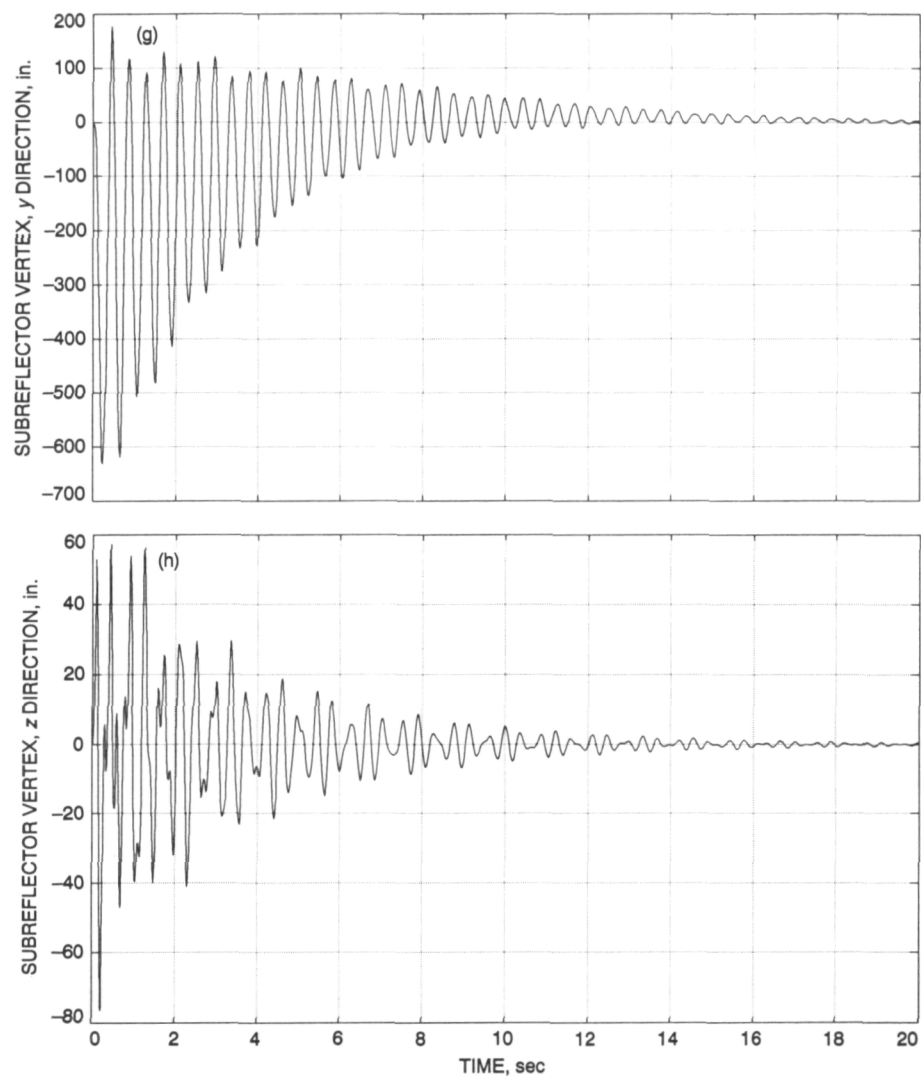


Fig. 14 (contd)

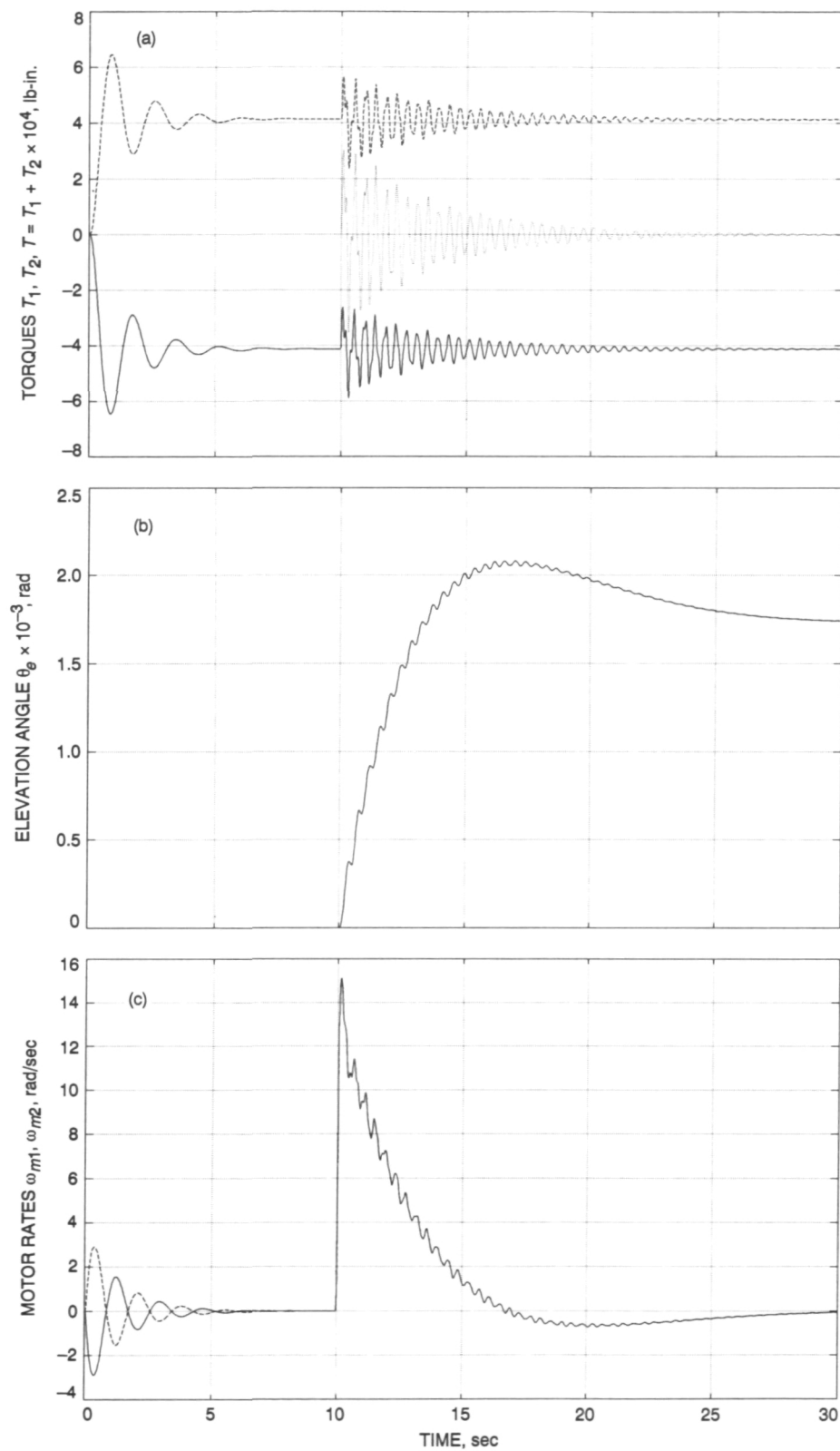


Fig. 15. Step responses of the closed-loop system to the command and bias inputs: step $v_{bias} = 2.5 \text{ V}$ at $t = 0$ and step $\theta_o = 0.1 \text{ deg}$ at $t = 10 \text{ sec}$.

Appendix A

Parameters of the DSS 13 Antenna

I. Structural Parameters

The structural parameters for the DSS 13 antenna are:

$$\alpha = 28 \text{ deg}$$

$$\zeta_i = 0.005, \quad i = 1, \dots, 10$$

$$\Omega = \text{diag}(0, 12.4482, 12.5833, 13.3612, 19.7028, 20.1747, 25.2584, 25.8547, 27.9621, 28.7060)$$

$$M_m = \text{diag}(215.4239, 223.8522, 38.1610, 17.1415, 243.8451, 58.5778, 95.1770, 804.4224, 18.4959, 18.3158)$$

Node numbers: $nc = 5380, no = 86302, nu = 86881$,
and $nb = 41212$

Modal matrix: $\Phi = [\phi_1, \phi_2, \dots, \phi_{10}]$

$\phi_i = i$ th mode shape, ϕ_i

$$= [\phi_{i1}, \phi_{i2}, \dots, \phi_{i12}]^T$$

where:

$\phi_{i1} = i$ th mode component, node nc, x direction

$\phi_{i2} = i$ th mode component, node nc, y direction

$\phi_{i3} = i$ th mode component, node nc, z direction

$\phi_{i4} = i$ th mode component, node no, x direction

$\phi_{i5} = i$ th mode component, node no, y direction

$\phi_{i6} = i$ th mode component, node no, z direction

$\phi_{i7} = i$ th mode component, node nu, x direction

$\phi_{i8} = i$ th mode component, node nu, y direction

$\phi_{i9} = i$ th mode component, node nu, z direction

$\phi_{i10} = i$ th mode component, node nb, x direction

$\phi_{i11} = i$ th mode component, node nb, y direction

$\phi_{i12} = i$ th mode component, node nb, z direction

The modes (in MATLAB notation):

$$\begin{aligned} \phi_1^T &= [0.00000012, \quad -0.00000060, \quad 0.00000003, \dots \quad \% \text{ (node } nc, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad 0.17696733, \quad 0.34187825, \dots \quad \% \text{ (node } no, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad -0.00000005, \quad -0.00000010, \dots \quad \% \text{ (node } nu, x, y, z \text{ components)} \\ &\quad 0.00000012, \quad -0.04215388, \quad 0.00000003]; \quad \% \text{ (node } nb, x, y, z \text{ components)} \\ \phi_2^T &= [0.16089562, \quad -0.04828905, \quad -0.01058891, \dots \quad \% \text{ (node } nc, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad -0.00320542, \quad 0.00104308, \dots \quad \% \text{ (node } no, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad 0.00012084, \quad 0.00023455, \dots \quad \% \text{ (node } nu, x, y, z \text{ components)} \\ &\quad 0.18123627, \quad -0.05316486, \quad -0.01040931]; \quad \% \text{ (node } nb, x, y, z \text{ components)} \\ \phi_3^T &= [-0.02516621, \quad -0.07031944, \quad 0.00690156, \dots \quad \% \text{ (node } nc, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad -0.00103347, \quad 0.00003283, \dots \quad \% \text{ (node } no, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad -0.00024008, \quad -0.00046601, \dots \quad \% \text{ (node } nu, x, y, z \text{ components)} \\ &\quad -0.02842479, \quad -0.07567856, \quad 0.00688137]; \quad \% \text{ (node } nb, x, y, z \text{ components)} \\ \phi_4^T &= [0.00351127, \quad 0.03635180, \quad -0.00294988, \dots \quad \% \text{ (node } nc, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad 0.00089782, \quad -0.00009626, \dots \quad \% \text{ (node } no, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad 0.00012046, \quad 0.00023381, \dots \quad \% \text{ (node } nu, x, y, z \text{ components)} \\ &\quad 0.00397642, \quad 0.03904565, \quad -0.00295351]; \quad \% \text{ (node } nb, x, y, z \text{ components)} \\ \phi_5^T &= [0.01726125, \quad 0.21422039, \quad -0.02128878, \dots \quad \% \text{ (node } nc, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad 0.41745298, \quad -0.03668607, \dots \quad \% \text{ (node } no, x, y, z \text{ components)} \\ &\quad 0.00000000, \quad 0.01317732, \quad 0.02557758, \dots \quad \% \text{ (node } nu, x, y, z \text{ components)} \\ &\quad 0.01343854, \quad 0.25263344, \quad -0.02139128]; \quad \% \text{ (node } nb, x, y, z \text{ components)} \end{aligned}$$

$$\begin{aligned}
\phi_6^T &= \begin{bmatrix} 0.08392051, & 0.00028101, & -0.03029582, \dots \\ 0.00000000, & -0.00560731, & -0.00120284, \dots \\ 0.00000000, & -0.00020205, & -0.00039219, \dots \\ 0.08607231, & -0.00154962, & -0.03083227; \end{bmatrix} \begin{matrix} \% \text{ (node } nc, x, y, z \text{ components)} \\ \% \text{ (node } no, x, y, z \text{ components)} \\ \% \text{ (node } nu, x, y, z \text{ components)} \\ \% \text{ (node } nb, x, y, z \text{ components)} \end{matrix} \\
\phi_7^T &= \begin{bmatrix} -0.02218441, & 0.16344080, & -0.01364561, \dots \\ 0.00000000, & -0.02133899, & 0.00544590, \dots \\ 0.00000000, & 0.00065982, & 0.00128072, \dots \\ -0.01743287, & 0.15653900, & -0.01357982; \end{bmatrix} \begin{matrix} \% \text{ (node } nc, x, y, z \text{ components)} \\ \% \text{ (node } no, x, y, z \text{ components)} \\ \% \text{ (node } nu, x, y, z \text{ components)} \\ \% \text{ (node } nb, x, y, z \text{ components)} \end{matrix} \\
\phi_8^T &= \begin{bmatrix} -0.30854177, & 0.04674442, & 0.18580163, \dots \\ 0.00000000, & 0.14007292, & 0.79574209, \dots \\ 0.00000000, & 0.05702746, & 0.11069132, \dots \\ -0.23322315, & 0.05969981, & 0.18902462; \end{bmatrix} \begin{matrix} \% \text{ (node } nc, x, y, z \text{ components)} \\ \% \text{ (node } no, x, y, z \text{ components)} \\ \% \text{ (node } nu, x, y, z \text{ components)} \\ \% \text{ (node } nb, x, y, z \text{ components)} \end{matrix} \\
\phi_9^T &= \begin{bmatrix} -0.00970566, & -0.01171182, & 0.00170980, \dots \\ 0.00000000, & -0.02275614, & 0.02106441, \dots \\ 0.00000000, & -0.00178816, & -0.00347085, \dots \\ -0.01092381, & -0.01958129, & 0.00171795; \end{bmatrix} \begin{matrix} \% \text{ (node } nc, x, y, z \text{ components)} \\ \% \text{ (node } no, x, y, z \text{ components)} \\ \% \text{ (node } nu, x, y, z \text{ components)} \\ \% \text{ (node } nb, x, y, z \text{ components)} \end{matrix} \\
\phi_{10}^T &= \begin{bmatrix} -0.05660426, & 0.01088616, & 0.01266761, \dots \\ 0.00000000, & 0.00316858, & 0.00525204, \dots \\ 0.00000000, & 0.00090903, & 0.00176443, \dots \\ -0.05696018, & 0.00916372, & 0.01287218; \end{bmatrix} \begin{matrix} \% \text{ (node } nc, x, y, z \text{ components)} \\ \% \text{ (node } no, x, y, z \text{ components)} \\ \% \text{ (node } nu, x, y, z \text{ components)} \\ \% \text{ (node } nb, x, y, z \text{ components)} \end{matrix}
\end{aligned}$$

II. Rate-Loop System Parameters

The rate-loop system parameters for the DSS 13 antenna are:

$$\begin{aligned}
k_1 &= 716.197 \text{ V sec/rad} \\
k_m &= 15.72 \text{ lb/A} \\
k_b &= 1.79 \text{ V sec/rad} \\
k_s &= 0.8 \text{ V/V} \\
k_{tach} &= 0.0384123 \text{ V sec/rad} \\
k_r &= 80 \text{ V/sec V, range: 49 - 83} \\
k_i &= 87.13 \text{ V/sec V} \\
k_{ii} &= 0.1 \\
k_{pp} &= 0.5 \\
k_{cur} &= 0.12658 \text{ V/A} \\
k_f &= 54 \text{ V/V} \\
k_{go} &= 1.5 \times 10^7 \text{ lb/rad}
\end{aligned}$$

$$\begin{aligned}
k_{ctfr} &= 0.33 \\
k_{ct} &= 0.11111 \\
k_{bs} &= 0.66 \\
\tau_1 &= 0.0063662 \text{ sec} \\
\tau_2 &= 0.094 \text{ sec} \\
\tau_3 &= 0.002 \text{ sec} \\
\tau_4 &= 0.00484 \text{ sec} \\
\tau_5 &= 0.0021 \text{ sec} \\
\tau_6 &= 0.7304 \text{ sec} \\
N &= 354 \\
J_m &= 1.236 \text{ lb/sec}^2 \\
R_a &= 0.456 \Omega \\
L_a &= 0.011 \text{ H}
\end{aligned}$$

Appendix B

From the Finite-Element Model to the State-Space Model

The state-space model of a flexible structure is obtained from its finite-element model, which consists of the mass M ($m \times m$), stiffness K ($m \times m$), input B_o ($m \times s$), output C_{oq} ($r \times m$), C_{ov} ($r \times m$) matrices, the input $u(t)$ ($s \times 1$), and output $y(t)$ ($r \times 1$). The input-output relationship is given by the second-order differential equation

$$M \ddot{q} + Kq = B_o u, \quad y = C_{oq} q + C_{ov} \dot{q} \quad (\text{B-1})$$

where q is the vector of structural displacements.

Consider now a modal matrix Φ ($m \times p$), $p \leq m$, which consists of p eigenvectors ϕ_i (mode shapes), $i = 1, \dots, p$:

$$\Phi = [\phi_1, \phi_2, \dots, \phi_p] \quad (\text{B-2})$$

which diagonalize M and K :

$$M_m = \Phi^T M \Phi, \quad K_m = \Phi^T K \Phi \quad (\text{B-3})$$

i.e., M_m and K_m are diagonal ($p \times p$) matrices of modal mass and stiffness. When a new variable q_m ($p \times 1$) is introduced such that

$$q = \Phi q_m \quad (\text{B-4})$$

and left-multiplying Eq. (B-1) by Φ^T , one obtains either

$$\Phi^T M \Phi \ddot{q}_m + \Phi^T K \Phi q_m = \Phi^T B_o u, \quad (\text{B-5a})$$

$$y = C_{oq} \Phi q_m + C_{ov} \Phi \dot{q}_m$$

or

$$M_m \ddot{q}_m + K_m q_m = \Phi^T B_o u, \quad y = C_{oq} \Phi q_m + C_{ov} \Phi \dot{q}_m \quad (\text{B-5b})$$

or

$$\ddot{q}_m + M_m^{-1} K_m q_m = M_m^{-1} \Phi^T B_o u, \quad (\text{B-5c})$$

$$y = C_{oq} \Phi q_m + C_{ov} \Phi \dot{q}_m$$

Denote $M_m^{-1} K_m = \Omega^2$, where Ω is a diagonal ($p \times p$) matrix of natural frequencies (rad/sec). At this stage a damping matrix Z is introduced, $Z = \text{diag}(\zeta_i)$, $i = 1, \dots, p$, such that $2Z\Omega = M_m^{-1} D_m$, and D_m is a modal damping matrix (assumed to be known), so that from Eq. (B-5c) the modal model is acquired:

$$\ddot{q}_m + 2Z\Omega \dot{q}_m + \Omega^2 q_m = M_m^{-1} \Phi^T B_o u, \quad (\text{B-6})$$

$$y = C_{oq} \Phi q_m + C_{ov} \Phi \dot{q}_m$$

Define a state variable x

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} q_m \\ \dot{q}_m \end{bmatrix} \quad (\text{B-7})$$

then, Eq. (B-6) can be presented as a set of first-order equations:

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -\Omega^2 x_1 - 2Z\Omega x_2 + M_m^{-1} \Phi^T B_o u$$

$$y = C_{oq} \Phi x_1 + C_{ov} \Phi x_2$$

or in the following form

$$\dot{x} = Ax + Bu, \quad y = Cx \quad (\text{B-8a})$$

where

$$A = \begin{bmatrix} 0 & I \\ -\Omega^2 & -2Z\Omega \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ M_m^{-1} \Phi^T B_o \end{bmatrix}, \quad (\text{B-8b})$$

$$C = [C_{oq} \Phi \quad C_{ov} \Phi]$$

is the sought state-space model in modal coordinates.

Appendix C

A Balanced Representation of a Linear System

Consider a stable, linear, time-invariant system:

$$\dot{x} = Ax + Bu, \quad y = Cx, \quad x(0) = x_o \quad (\text{C-1})$$

where $x \in R^n$ is the system state, $u \in R^p$ is the system input, $y \in R^m$ is the system output, and (A, B, C) is the system state-space representation. The system controllability and observability grammians W_c and W_o are solutions of the Lyapunov equations

$$AW_c + W_cA^T + BB^T = 0, \quad A^TW_o + W_oA + CC^T = 0 \quad (\text{C-2})$$

The system representation is balanced if its controllability and observability grammians are diagonal and equal. Hence, for the balanced representation $(A_b, B_b, C_b) = (T^{-1}AT, T^{-1}B, CT)$, the following is true:

$$W_c = W_o = \Gamma^2, \quad \Gamma = \text{diag}(\gamma_1, \dots, \gamma_n), \quad \gamma_i \geq 0, \quad i = 1, \dots, n \quad (\text{C-3})$$

where T is a linear transformation and γ_i is the i th Hankel singular value of the system.

The transformation T is determined as follows:

$$T = P\Gamma^{-1}, \quad T^{-1} = \Gamma^{-1}V^TQ \quad (\text{C-4})$$

The matrices Γ , V , and U are obtained from the singular value decomposition of the matrix H :

$$H = V\Gamma^2U^T \quad (\text{C-5})$$

where $H = QP$, and P and Q form the following decomposition of grammians:

$$W_c = PP^T, \quad W_o = Q^TQ \quad (\text{C-6})$$

for example, Cholesky, or singular-value decomposition.

1994025120

N94-29623

TDA Progress Report 42-105

442614

May 15, 1991

Radio Science Ground Data System for the Voyager-Neptune Encounter, Part I

E. R. Kursinski

Tracking Systems and Applications Section

S. W. Asmar

Telecommunications Systems Section

The Voyager radio science experiments at Neptune required the creation of a ground data system array that includes a Deep Space Network complex, the Parkes Radio Observatory, and the Usuda deep space tracking station. The performance requirements were based on experience with the previous Voyager encounters, as well as the scientific goals at Neptune. The requirements were stricter than those of the Uranus encounter because of the need to avoid the phase-stability problems experienced during that encounter and because the spacecraft flyby was faster and closer to the planet than previous encounters. The primary requirement on the instrument was to recover the phase and amplitude of the S- and X-band (2.3 and 8.4 GHz) signals under the dynamic conditions encountered during the occultations. The primary receiver type for the measurements was open loop with high phase-noise and frequency stability performance. The receiver filter bandwidth was predetermined based on the spacecraft's trajectory and frequency uncertainties.

I. Introduction

This is the first of two articles on the Deep Space Network (DSN) implementation in support of the radio science activities of the Voyager-Neptune encounter. This article will discuss the key requirements for the implementation, but not the actual design and implementation efforts. The specifications discussed here are a combination of the *Voyager Science Instrumentation Requirements Document* (SIRD) and the DSN's requirements and goals. This arti-

cle is intended as a tutorial, creating in the process a context in which to better present the various requirements and, subsequently, the design and implementation efforts. This article may also be useful for deriving radio science experiment requirements for future missions.

Neptune and its satellites represent the fourth planetary system encountered by Voyager and at least the fourth upgrade to the DSN radio science system since its initial

implementation. Consequently, the development of the requirements incorporated the considerable knowledge and understanding accumulated during past encounters. Nevertheless, the process of choosing a level of performance to be a requirement is always a bit tricky because the desired performance for the radio science instrument is open-ended; that is, the better the instrument performs, the better the scientific results will be. The scientific goals and the technology necessary to achieve them must be combined to set the actual, achievable performance. "Requirements," of course, also include a third aspect, namely, commitment, which implies a function or level of performance that is virtually guaranteed.

The Voyager radio science experiments at Neptune fell into two fundamental classes. The first class was composed of measurements related to gravity and celestial mechanics, where the gravity field is inferred from the motion of the spacecraft measured with the range and range rate using the telecommunications signals, as well as very long baseline interferometry (VLBI) and optical navigation data. This article focuses on the second class, which characterizes the electromagnetic properties of the media through which the signals pass by measuring the effects they induce on the received signal amplitude and phase. "Occultation" refers to the interruption of the light from a celestial body or of the signals from a spacecraft by the intervention of a celestial body. Occultation is also used when the media of interest fall along the line of sight between the Earth and a spacecraft. The occultation measurements form a special class of remote-sensing experiments that are unique because of the coherent nature of the probing signal. For example, in the case of stellar occultations, the signal is broadband and incoherent and, therefore, only its intensity can be used to recover information about the media being probed. In the case where a coherent signal source is used, both phase and amplitude information are available.

The radio science instrument is unique as a Voyager instrument because the ground tracking stations are part (in fact the larger part) of the instrument itself. This arrangement provides the opportunity to improve the instrument after launch. In the case of Voyager, experiments are performed using the downlink signal from the spacecraft to the Earth. This configuration has been used because it requires the least equipment on board the spacecraft. A proposal has been made for future spacecraft to place a receiver and data processor on board the spacecraft to take advantage of the 20–30 dB improvement in the signal-to-noise ratio (SNR) of the uplink-versus-downlink telecommunication signals.

II. Neptune Occultation Experiment Overview

The Voyager trajectory was chosen so that, during the course of its flyby, the spacecraft, as viewed from Earth, would be occulted by the atmosphere, ionosphere, and rings of Neptune, as well as by the atmosphere and ionosphere of Triton. In a very real sense, there were five different occultation observations. That number can, in fact, be doubled to ten because, in each case, an occultation occurs on either side of the planet or moon. A brief description of each of these different measurements follows.

A. Neutral Atmospheres

In an atmospheric occultation experiment, the S- and X-band signals from the spacecraft pass through the planetary atmosphere on their way to Earth. As they travel, they are slowed by refraction, which causes their paths to bend and increase in length relative to their length in the absence of the planet's atmosphere. The increased optical path length is essentially determined by using the signal phase as a gauge. Knowledge of the relative positions of the spacecraft, occulting planet, and Earth, as well as the gravity field of the planet, allows the bending angle as a function of altitude in the planet's atmosphere to be reconstructed from the phase data. The phase data can be inverted to recover the refractivity of the atmosphere as a function of altitude by solving an Abelian integral equation [1]. The number density of the atmosphere can be derived from the refractivity by assuming knowledge of the atmospheric constituents derived from other instruments and atmospheric models. The pressure can be derived by vertically integrating the number density multiplied by the weight of an average molecule by using the equation of hydrostatic equilibrium in which atmospheric pressure balances the weight of the atmosphere. Given both the number density and pressure as functions of altitude, the temperature versus altitude is then derived from the equation of state (the ideal gas law in the simplest case).

This process begins with the ray above the top of the detectable atmosphere. As the periapsis of the ray descends into the atmosphere, the bending-angle profile is constructed as a function of altitude, one layer at a time, until the maximum depth probed by the beam is reached. The outbound occultation is processed in time-reversed order, again beginning with the ray at the top of the atmosphere and working backward, down into the atmosphere. Therefore, the inbound and outbound occultations are processed separately. If the signal is continuously visible throughout the entire two occultations, at least one additional constraint can be placed on the atmosphere so

that the two occultation inversions match at the deepest point where their measurements overlap. This technique was used during the Uranus encounter to estimate the wind velocity in this deeper region of the planet's atmosphere [2].

Although the key measured parameter is the phase, amplitude information is available as well. The nearly exponential change of refractivity with altitude causes differential bending across the vertical dimension of the radiation pattern from the spacecraft, which causes the ray to defocus and reduces the beam's energy flux density. This, in turn, causes the amplitude of the received signal to decrease dramatically. The same effect can be used to measure refractivity. In the case of a coherent microwave signal, however, the phase is a far more sensitive measure of refractivity. Consequently, the expected defocusing of the signal level is estimated from the phase data and subtracted from the measured signal amplitude. Any residual amplitude attenuation is used to estimate the density of any microwave-absorbing constituents that might be present in the atmosphere.

The Neptunian atmospheric occultation was designed to measure the pressure and temperature of two slices through the atmosphere, over a range from approximately 0.1 millibar to several bars. Triton was expected to have a very tenuous atmosphere (if any) of tens of microbars. It was expected that the extra path length in this case would be extremely small (subcycle) and would push the limit of the technique beyond that which had previously been accomplished. It is important to note that this technique measures increased path length over the duration of the occultation. Therefore, the knowledge of the absolute distances involved is needed to a much lower accuracy than the accuracy of the *change* in path length over the duration of the occultation. A typical duration for the various planetary atmospheres encountered by Voyager was 1,000 sec, whereas the Triton occultation was expected to last only a few seconds.

B. Ionospheres

In the case of ionospheres, the electron density integrated along the signal path causes a shift in the signal phase that is inversely proportional to the signal frequency. The linear combination, $S-3/11X$, of the phase of the coherently related S- and X-band signals can then be formed, which cancels the nondispersive effects and leaves a quantity that is proportional to the integrated electron density, commonly referred to as the total electron content (TEC). Starting with measurements above the detectable ionosphere, the number density of the ionosphere as a function of altitude above the planet can be estimated from the

differential phase data. The estimate assumes that the effect is large enough to stand out against the signatures of other dispersive effects due to instrumentation and intervening media.

C. Ring Occultations

In the case of an occultation by a ring of material, there is a modified coherent signal, as well as incoherent scattered signals, that can be received at the Earth. The amplitude and phase of the coherent signal can be used to measure the complex microwave opacity of planetary rings. The differential opacity and phase delays and the differential levels of forward scattering of the S- and X-band signals can be used to estimate the particle-size density and distribution.

Because of diffraction, the radial resolution of the opacity measurements might at first seem very coarse, as compared with the stellar occultation measurements made by the Voyager Photopolarimeter Subsystem (PPS) at infrared (IR) and visible wavelengths. Optical wavelengths are approximately 100,000 times smaller than the X-band wavelength, which makes the Fresnel scale (diffraction limit) size some 300 times larger at microwave wavelengths. However, because the microwave signals are coherent and the geometry is changing, the effects of diffraction can be largely removed in a manner similar to synthetic aperture radar, which results in a resolution of hundreds to tens of meters instead of tens of kilometers. This technique and its limitations are discussed in [3]. In addition, even with the tremendous distance between the Voyager spacecraft and the Earth, the radio science instrument has more than two orders of magnitude more dynamic range than the PPS instrument, which enables it to sense a wider range of opacities and more accurately characterize a range of opacities common to both instruments. These two opacity measurements by the PPS and the radio science instrument are, of course, complementary because of their vastly different wavelengths.

As previously mentioned, the difference between the levels of the approximately forward-scattered S- and X-band signals is used with the opacity measurements to infer the particle-size distribution and densities. The scattered signals are separated from the coherent signal by using the high-frequency resolution due to the coherence of the measurement system. In the case of Saturn's A ring, the large difference between the S- and X-band levels implied the presence of a significant number of centimeter-sized particles [4]. In the case of Uranus, the opacity at both wavelengths was large and virtually identical, which implied that the particle sizes were probably larger than either wavelength [5].

III. Key Requirements for the Instrument

The primary requirement for the instrument is to recover the phase and amplitude of the S- and X-band signals under the dynamic conditions encountered during the occultations. Data gaps are undesirable because they disrupt the continuous phase record or gauge. This happens because the spatial resolution of the key parameters depends on the signal phase measured over some finite range of time. The specifications were written so that 100-percent recovery of the data during the critical periods is highly probable.

An attempt was made to make the measurements as sensitive as possible. The fundamental performance limitations of the instrument were identified, and a goal was set to make all other sources of error as small as possible, given the constraints of available and reliable technology. As it became apparent that this goal was achievable, a set of requirements was established.

Finally, an attempt was made to enable the equipment installed at the DSN sites to meet their long-term needs. This meant improving in the ease of system operation, as well as addressing, where appropriate, the needs of future Galileo and Mars Observer experiments.

IV. Instrument Conceptual Design

As mentioned above, the instrument is made up of the spacecraft- and Earth-based microwave radio systems, which were created primarily to support telecommunications and navigation. The mode generally used for occultations is the so-called one-way Doppler mode, which is also referred to as bistatic radar. In this configuration, the S- and X-band signals radiating from the spacecraft are generated from a reference oscillator on board the spacecraft. These signals are received at the Earth and are measured there against a reference oscillator. This process differs from the "two-way" mode, where an uplink signal is generated on the ground and transmitted to the spacecraft. The spacecraft then receives the signal and uses it as a reference frequency to generate a downlink signal that is radiated back to Earth.

The advantages of using the one-way mode for an occultation experiment are two-fold. As the uplink signal passes through the intervening media before being received by the spacecraft, the spacecraft receiver (transponder) must phase lock onto the signal in order to use it as a reference from which to generate a downlink signal (which then passes through the occulting media a second time). The first problem is that the signal dynamics introduced by

the occulting media make it difficult for the transponder to acquire and maintain phase lock. On the outbound atmospheric occultation, the spacecraft receiver may take up to several minutes to lock onto the uplink signal, thereby losing a significant portion of the medium under study. The second problem is that the finite bandwidth of the transponder's transfer function will filter out some of the effects of interest. This configuration does place relatively stringent requirements on the phase stability of the oscillator on board the spacecraft. The ultrastable oscillator (USO) on board the two Voyager spacecraft is the one piece of equipment specifically placed there to support radio science. For such spacecraft as Pioneer Venus, the on-board oscillator does not have sufficient stability, and the inbound atmospheric occultations are generally performed in the two-way mode. The outbound occultations are not performed in this mode because of the time necessary for the transponder to lock onto the uplink signal.

The spacecraft side of the instrument, although basically fixed, has some flexibility. The transmitting levels of the S- and X-band signals are selectable within the constraints of the available power. With Voyager, power constraints make it possible for only one of the two signals to be at high power. Another important aspect of the transmitted power is that the telemetry phase modulation can be turned off for occultations, which puts all the transmitted power into the monochromatic carrier signals. (At X-band, this meant an increase of approximately 12.5 dB in the carrier level given the planned data rates for Neptune.) Another critical function that the spacecraft must perform is an extremely important maneuver during the Neptune atmospheric occultation, where the spacecraft's high-gain antenna is pointed at the expected image of the Earth as viewed through the atmosphere. The proper execution of this maneuver places very sharp constraints on the accuracy of the trajectory prediction, as well as on the attitude control system of the spacecraft.

The ground system requires a receiving antenna with a very high gain, implying a large dish antenna. This is followed by a low-noise amplifier (LNA), presumably cryogenically cooled for optimal performance, thus creating the front end of a receiving system with a very high gain to noise temperature (G/T) ratio. This is a feature of such telecommunication systems that is critical to the SNR achievable with the system upon which the achievable telemetry rates depend. The antennas used for telemetry can be shared and serve both purposes, which should be obvious because the radio science capability emerged from the telemetry and navigation systems for deep space probes.

Because the ground receiving equipment must handle the signal dynamics induced by the planetary media of interest, the primary receiver type for these measurements is open loop. "Open loop" means that there is no signal detection in real time. The entire signal spectrum of interest is captured in real time so that the signal phase and amplitude can be detected at some later time, which allows one iteratively to optimize parameters in the detection process. The lack of real-time detection occurs because it is so difficult to detect the dynamic signals in real time, and because the optimum detection algorithm may differ from one type of occultation to another. One particular problem involves multipath, where the signal can arrive simultaneously at the Earth via multiple paths. Because the signal is essentially monochromatic, the various paths the signal takes can be separated via the unique Doppler shift associated with each path. The narrowness of the spectral line limits the ability to distinguish among different paths.

The open-loop receiver heterodynes the signal at the antenna against a series of reference frequencies derived from a common, stable, reference signal to a much lower frequency signal where it can be digitized and recorded. The frequency of the resulting downconverted signal is the difference between the received signal frequency and that of the local oscillator (LO) chain and, therefore, this frequency contains the phase noise of the local oscillators, which makes the phase stability of these LO's a critical parameter in the instrument's sensitivity. To minimize the bandwidth needed to acquire the signal, one local oscillator is tuned along the predicted Doppler profile of the received signal, which removes most of the Doppler shift prior to recording the signal. The actual bandwidth requirements will be discussed further. Recovery of the microwave signal frequency and relative phase received at the antenna require that the frequency and phase of this tunable local oscillator be recorded.

The reference frequency at the Earth receiving station is used to generate both the reference frequencies from which the LO chain is derived and the clock used to time tag the received data. Because the LO chain and, therefore, the phase stability of the downconverted signal ultimately depend on the stability of the station reference, the stability over the occultation length is critical to these experiments.

In order to fully recover the dynamic range of the received signal spectrum and avoid any further degradation of the signal when subsequent copies of the data are generated, the downconverted signals are digitized prior to recording. The digitization adds a small amount of noise to the analog signal being recorded, but this can be made

negligible by using a sufficient number of bits and a stable sampling clock. The data are then recorded digitally on some form of bulk storage media (currently magnetic tapes) at the receiving site. Due to the data rates involved and the cost of transmitting the data via satellite to the Jet Propulsion Laboratory (JPL), the data are subsequently shipped to JPL within a couple weeks of the event. It is also very desirable that this media be compatible with the computer facilities of the investigators.

V. Filter Selection and Bandwidth Allocation

As discussed, one local oscillator in the open-loop receiver is tuned to reduce the required bandwidth. The narrowest open-loop filter bandwidth was selected to ensure that the Voyager signal would remain within the open-loop receiver bandwidth throughout the Neptune encounter. Figures 1 and 2 illustrate the expected Doppler and frequency rates during the encounter. The requirement was formulated after the completion of a study of the factors that contribute to frequency offsets and uncertainties at X-band.¹ There is a bandwidth requirement associated with each of the ring, atmospheric, and Triton occultation experiments at Neptune. The 3σ X-band requirements at Neptune are summarized in Table 1.

The selected filter bandwidth is the same for both frequency bands because the Doppler spread across the spacecraft antenna beamwidth is independent of frequency. The spacecraft antenna illuminates a region of the planet's rings, for example, and the resulting Doppler shift across that region results primarily from the spacecraft's motion. The Doppler shift for small angles is

$$F_{Doppler} \approx \frac{f\theta V}{c}$$

where F is the frequency of the straight-line signal to Earth, V is the component of the spacecraft velocity in the plane of the sky, c is the speed of light, and θ is the angle relative to the straight-line signal path to Earth. The maximum angle θ_{max} is limited by the beamwidth of the spacecraft antenna. Therefore, θ_{max} is inversely proportional to the frequency as

$$\theta_{max} \sim \frac{c}{FD}$$

¹ S. W. Asmar, *Signal Passband Frequency Allocation*, JPL Interoffice Memorandum Voyager-RSST-87-041 (internal document), Jet Propulsion Laboratory, Pasadena, California, October 21, 1987.

where D is the spacecraft antenna diameter, resulting in

$$F_{Doppler_{max}} \approx \frac{F\theta_{max}V}{c} = \frac{V}{D}$$

which is independent of signal frequency.

To reduce the frequency uncertainties, it was required that the receiver-tuning predictions be updated shortly before the encounter based on an improved spacecraft trajectory. Three late navigation solutions were critical for the limb-track maneuver and for determining the spacecraft arrival time. The arrival time is proportional to the frequency, and errors in it result in significant frequency errors due to the rapidly changing Doppler shift.

Factors that contributed an error greater than 1 kHz include the 1σ values of the errors listed below. These values were root-sum-squared and multiplied by 3 for the values in Table 1:

- (1) Timing error: The frequency offset due to a 5-sec timing error was estimated to be 1.3 kHz.
- (2) Radius error: The frequency offset due to a 50-km error in Neptune's radius was estimated to be 1.0 kHz.
- (3) Ring scattering: The frequency range due to scattering by the Neptunian rings was estimated to be 6 kHz at 3 dB off antenna Earth boresight and 10 kHz at 10 dB off boresight. Though Doppler is measured on a frequency scale, the antenna beamwidth requires the same bandwidth independent of the frequency band.
- (4) The Neptune Doppler uncertainty at Earth occultation ingress for an $(N-5)$ -day orbit determination (OD) cutoff date (latest estimate of late ephemeris update, LEU, see [13]) was 6.2 kHz, assuming radio-plus-optical narrow-angle information. The Neptune Doppler uncertainty at Earth occultation ingress for an $(N-2)$ -day OD cutoff date (latest estimate of late stored update, LSU) was 1.0 kHz, assuming radio-plus-optical narrow-angle information.
- (5) An arbitrary offset introduced to facilitate data post-processing of about 2.5 kHz was also taken into account.
- (6) The bandwidth requirement for the Triton occultation was estimated to be at least 8.4 kHz. Due to the tenuous nature of Triton's atmosphere, there was no requirement for modeling the atmosphere in the tuning predictions.

Factors below the 1-kHz contribution include:

- (1) Neptune's atmosphere: The uncertainties in frequency predictions in the modeling of the Neptunian atmosphere have been specified not to exceed 400 Hz by the Radio Science Team.
- (2) Ultrastable oscillator: Radiation effects may cause the USO frequency to shift by as much as 50 Hz.
- (3) DSP-R tuning error: The radio science predictions were expected to be generated with a frequency tuning error of less than 100 Hz. This took into account the predicted software requirements for other missions as well.

The conclusion of this study was to require the use of a 20-kHz filter as a minimum bandwidth. This was selected as a prime filter and a 45-kHz filter was required for backup. The prime filter would require a sampling rate of 50 kps, which was an existing capability of the system used for the Uranus encounter. The backup filter would require a sampling rate of 100 kps, accomplished by pairing up two analog-to-digital convertors per frequency band, each at a base-sampling rate of 50 kps.

VI. Implementation Purpose

The apparent maturity of the DSN radio science system and the generally successful support to the Uranus encounter leads to a question of whether an upgrade to the ground system was needed for the Neptune encounter. The two primary reasons for the upgrade were that new capabilities were needed in the array of ground antennas and that the requirements for Neptune were tighter than those specified for Uranus.

For the Neptune encounter, the desire was to increase the receiving aperture to offset the increased distance from Uranus to Neptune. This not only implied multiple receiving antennas, but also the possibility of non-DSN observatories and tracking stations. As will be discussed, the particular array of antennas chosen included DSS 43, the Parkes Radio Observatory in Australia, and the ISAS tracking station outside of Usuda, Japan. This implied a reimplementaion at Parkes (with a number of improvements), as well as an entirely new implementation in Japan.

The Neptune requirements were tighter than those of the Uranus encounter for two basic reasons. The first reason was the closer and faster flyby of Neptune, which increased the Doppler rates and the requirements on the

support system's ability to provide prediction updates very close to the time of closest approach. The second requirement was based on tightened performance specifications, particularly in the area of phase stability, to avoid problems experienced during the Uranus encounter. A serious problem occurred at Parkes when the phase stability of an LO seriously degraded the Parkes data's ultimate utility. Another dramatic problem was the very late deletion of the DSS 42 X-band receiver channel from the Uranus radio science array due to problems of coupling between the DSS 42 and 43 IF signals within the receiver. Better specifications for the necessary system performance and tighter performance requirements resulted from these problems as well as the justification of the implementation of special equipment designed to test the system at the level of those requirements. Very limited capabilities had previously been provided and their development, along with a testable set of performance specifications, played a major role in the ultimate success of the implementation.

Reliability had always been a concern of these implementations and this encounter was no exception. Because this encounter was probably the only time that Neptune would be viewed up close for another 20 years, system reliability was considered to be crucial. The stated Voyager SIRD specification was 99 percent probability of 100 percent data collection over any four-hour time span. That specification was not formally accepted by the DSN because of the virtually untestable nature of the requirement for at least 500 such tests with only a single allowable failure. Even though the formal requirement was not accepted, the purpose of the requirement was understood and led to a design with virtually 100-percent redundancy throughout the entire ground array.

In addition, from a longer term DSN perspective, there was a strong desire to improve the operability of the system for multimission support, which would ultimately lead to better the system reliability and human operator interaction. That desire led to a series of improvements in the monitor and control aspects of the system.

VII. Arraying Selection and Expected Performance

The purpose of using an array of ground antennas for the reception of the radio occultation data was three-fold. The first was to improve the SNR of the received S- and X-band signals. A major fraction of the SNR reduction relative to Uranus was compensated for by the increased aperture and efficiency of the DSN 70-m antennas. The second purpose was to provide redundancy so that if an

antenna failed, other antennas would continue to acquire critical data. The third purpose was to separate effects of interest from those due to instruments. Multiple receiving sites would allow isolation of effects unique to a particular site.

The array for Uranus differs from the Neptunian telemetry array in two ways. First, the arraying is done at microwave frequencies, as opposed to video-band modulation signals. Second, the signals are combined after the acquisition is complete, as opposed to the real-time combining by the telemetry array. Real-time combining was due primarily to the dynamic nature of these signals, which made near optimal real-time combining virtually impossible. There are, therefore, no requirements for high-rate real-time data communication links, which is important because the data rates are on the order of 800 kbps. Playback of certain portions of the data was required for quick-look analysis immediately following the encounter.

In selecting the antennas nominally desired to take part in the array, the following factors were considered:

- (1) The expected use of at least one DSN 70-m antenna
- (2) The 5.5-hour separation between the Neptune and Triton events
- (3) The desire for reception at high elevation angles to reduce noise temperature and the amount of Earth atmosphere and ionosphere through which the signals had to pass
- (4) The desire to utilize the SNR improvements made for the telemetry array
- (5) The desire to use a few large antennas to reduce complexity and level of effort
- (6) The capability to array both S- and X-band signals

Given the desire for reception at high elevation angles and the -22 -deg declination of Neptune at the time of encounter, the most desirable DSN 70-m antenna was DSS 43. As shown in Fig. 3, the entire set of events at Neptune and Triton could be received at this one location at relatively high elevation angles near 45 deg. Such reception was also desirable because the gain of the DSN 70-m antennas peaks in this range, which also allows use of the Parkes radio telescope—already a part of the telemetry array with DSS 43 and with approximately the same elevation angle coverage as DSS 43. Parkes had been outfitted as an X-band-only receive site for the Uranus encounter and was expected to be in the same configuration for the Neptune encounter. Parkes, with the lack of S-band telemetry-arraying requirements, combined with

its small cage at the prime focus of the antenna housing the LNA's and first downconversion stage of the receiver, was expected to fulfill the minimum needs of the three arraying functions at X-band, but not of those at S-band.

Japan is located approximately at the same longitude as Parkes, and the Japanese space agency, ISAS, had built a 64-m tracking station in Usuda, Japan, to provide S-band tracking of its two spacecraft, Suisei and Sakegake, which encountered Halley's comet in 1986. With the use of a cryogenically cooled S-band LNA, the Usuda antenna had demonstrated a G/T comparable to a DSN 64-m antenna [6] and had been used in the past both to track ICE, a National Aeronautics and Space Administration (NASA) satellite, and support the space VLBI experiment with the Tracking and Data Relay Satellite System (TDRSS). Usuda's ability to track at S-band potentially fulfilled the S-band arraying needs with a G/T nominally within 1 dB of DSS 43. It had the disadvantage of being located at a 36 deg north latitude, which caused the encounter events to occur at approximately a 20-deg elevation, but both the Neptune and Triton occultation events would be visible from a single site. Given these considerations, the nominal array was defined as DSS 43 operating at both S- and X-band, Parkes operating at X-band, and Usuda operating at S-band. The desire for both Parkes and Usuda to operate at both S- and X-band was also stated.

The expected array performance relative to Uranus is summarized in Table 2. The G/T for DSS 43 was to be improved over that at Uranus by 1.9 dB at X-band and by perhaps 1.4 dB at S-band, with the extension of the antenna to 70 m in diameter. The Parkes radio telescope performance was to be the same as that at Uranus, which meant that its G/T would be about 3 dB less than that of DSS 43. The Usuda S-band performance was expected to be within about 1 dB of that of DSS 43. Usuda has less area but a higher efficiency and a lower zenith noise temperature, and also a lower elevation angle because of its northern latitude location. Table 2 also includes the approximately 2-dB loss in received X-band signal level at DSS 43 due to a mispositioned subreflector during the Uranus encounter.

The lower S-band noise temperature at Usuda was partially due to the assumption that DSS 43 would have its diplexer in place to allow uplink operations during the encounter pass. As it turned out, because of clever planning by the Voyager Radio Science Team, uplink was not required from DSS 43 and the diplexer was not used, which lowered the system noise temperature (SNT) from 20 to 16 K, an improvement of 1 dB in SNR. Removing the

diplexer also reduced Usuda's SNR to about 2 dB below that of DSS 43 and the improvement at S-band over that of DSS 43 alone to about 2.1 dB, or a 21-percent reduction in root-mean-square (rms) amplitude and phase jitter.

The importance of the redundancy provided by such an array cannot be overstated. In the event of a failure of any of these three antennas, a dual frequency data set would continue to be collected at the other two sites. DSS 43, of course, would be the greatest loss because of its higher SNR and dual frequency capability. However, Parkes and Usuda would essentially combine to produce individual S- and X-band data sets equivalent to those acquired at a DSN 64-m station. In addition, either DSS 45 or 42 could be used with a reduced SNR, particularly to produce the dual frequency data sets for characterizing the ionospheres of Neptune and Triton. The redundancy also made the data-acquisition process less susceptible to local weather effects, a serious concern during August in Australia.

VIII. General Stability Discussion

The SNR, as well as the phase and amplitude stabilities of the received signal, is extremely important to the sensitivity and accuracy of the parameters ultimately extracted from the received signal phase and amplitude. The effect of the SNR is to add a random phaser to the signal phaser modulating the phase and amplitude of the signal. For SNRs greater than approximately 10, the total noise power of the random phaser is statistically equally distributed between the amplitude and phase, which makes the amplitude and phase noise power spectral densities each half of the total power spectral density. The thermal noise associated with the SNR is also white or uncorrelated from sample to sample, assuming that the measurement noise bandwidth is the reciprocal of the sample interval. The implications for the different types of occultation measurements are briefly discussed below.

A. Atmosphere

The SNR is a dominant source of error in the estimates of temperature and pressure of the upper regions of the atmosphere, which was very important in the case of Triton because of the tenuous nature of its atmosphere. The SNR is also important deep in the atmosphere because of defocusing, which causes a loss in the signal level, typically at least 20 dB for the Voyager occultations. The actual defocusing effects on the intensity scales as a parameter M , which for an isothermal atmosphere equals $[1 + (\alpha D)/H]^{-1}$ where α is the bending angle in radians,

D is the distance from the spacecraft to the limb, and H is the refractivity scale height [7]. At the deepest point in the Neptune occultation, the expected values for α , D , and H were approximately 15 deg, 40,000 km, and 25 km, respectively, which resulted in an expected intensity reduction of 400 or 26 dB. The need for an even higher SNR exists when there are small-scale, vertical structures in the atmosphere, such as cloud decks, with small refractivity scale heights, where refractivity changes more rapidly with altitude than the background atmosphere. Such was the case with Uranus and was anticipated (correctly so) at Neptune. The scale height in these regions can be a factor of 10 smaller than the background atmosphere, which increases the defocusing in the region by a factor of 10 or more. Signal multipath is typically associated with these regions as well. Each of the multipath signals passes through a unique vertical region of the atmosphere and must, therefore, be recovered to characterize the atmosphere at and below this depth.

The SNR necessary for tracking signals passing deep into the atmosphere can be estimated as the minimum SNR needed for signal detection after accounting for the expected signal level loss due to atmospheric defocusing and absorption. The minimum SNR must also account for the noise bandwidth equivalent to the signal detection-integration interval. The minimum SNR needed for detection is typically taken to be 10 dB. The maximum integration time for detection is generally taken as the time for the signal to descend a Fresnel scale, F , the diffraction-limited vertical size of the geometric ray. In the upper atmosphere, where bending is insignificant, the Fresnel scale is $\sqrt{\lambda D}$, where λ is the wavelength and D is the distance from the spacecraft to the planet's limb. This descent time is the Fresnel scale divided by the component of the spacecraft velocity in the direction of local vertical at the deepest point of the ray in the atmosphere. In the case of Voyager, this leads to approximate X- and S-band integration times of 0.08 and 0.16 sec, respectively, at the top of the atmosphere. However, deep in the atmosphere, both the vertical dimension of the Fresnel scale and the vertical descent rate of the ray decrease due to bending. The vertical dimension of F scales as \sqrt{M} , whereas the velocity scales as M , which causes the integration time to scale as $1/\sqrt{M}$ and the associated noise bandwidth as \sqrt{M} . The minimum free-space SNR can then be estimated as

$$SNR_{0 FS} = \frac{10V_v\sqrt{M}}{FM} = \frac{10V_v}{F\sqrt{M}}$$

where $SNR_{0 FS}$ is the free-space signal-to-noise-spectral density ratio, V_v is the vertical component of the space-

craft velocity relative to the planet, and F is the geometric Fresnel scale at the limb of the planet in the absence of atmosphere.

The factor of 10 in the numerator is the minimum SNR required for reliable signal detection. In the case where $M = 1/400$, $SNR_{0 FS}$ should be at least 34 and 31 dB for X- and S-band, respectively.

Another item of interest is the characterization of absorbing material present in the atmosphere. Ammonia, a very common solar system constituent, absorbs more at X-band wavelengths than at S-band, which causes one to consider selecting the highest available transmitter power at S-band if the presence of ammonia is indeed anticipated. Ammonia was not expected in the Neptunian atmosphere at the depths that were probed by the radio signal, based on the Uranus experience, and the X-band transmitter in the high-power mode was selected as the optimum configuration. This assumption, in fact, proved to be incorrect, and ammonia was indeed present in sufficient quantities to cause the S-band signal to be detectable deeper in the atmosphere than the X-band signal despite an ~15-dB higher X-band free-space SNR.

Amplitude stability over time scales long enough to allow effects other than the SNR to become significant is also important for characterizing microwave-absorbing material in the atmosphere. The sensitivity to and accuracy of recovered absorption constituent number densities depend on instrumental amplitude stability. Typically, one looks for differential absorption between the S- and X-band signal levels because the absorption strongly depends on wavelength. Knowledge of the actual pointing of ground and spacecraft antennas during the observations becomes very important because the antennas are an obvious source of relative changes between S- and X-band signal levels in as much as antenna beamwidth scales with wavelength.

Instrumental phase stability sets limits on the accuracy of the recovered atmospheric temperature and pressure profiles. In the extreme case, the phase can be sufficiently unstable that the atmosphere cannot be separated from the phase noise. The transmitted signal frequency and phase during an occultation are unknown and must be estimated from the frequency measured just prior to the occultation interval. Given that the phase error in this estimate does not increase too rapidly over the time scale of an occultation, the phase uncertainty sets the upper altitude limit at which a specific temperature accuracy can be achieved. As the ray descends deeper into the atmosphere, the temperature error tends to decrease because

the changing path length increases exponentially with decreasing altitude and generally much faster than the uncertainty in the phase change of the system, particularly of the oscillators involved. A relevant time scale is the time it takes the periapsis of the ray to descend one scale height in the atmosphere. A scale height is the altitude change over which the atmospheric parameter, such as density or pressure, changes by e . In the case of Neptune, with an approximate scale height of 25 km, that time was about 3 to 4 sec in the upper atmosphere.

B. Ionosphere

The ionosphere is characterized by measurements of $\phi_S - 3/11\phi_X$, the differential phase between S- and X-band signals. The effect of SNR on this measurement is dominated by the S-band SNR because of the combination of the 3/11 scaling factor for the X-band and the 15-dB higher X-band SNR. As shown in the following equation, the S-band SNR makes the S-band responsible for more than 99 percent of the high Fourier frequency noise:

$$\begin{aligned}\sigma_{\Delta\phi}^2 &= \sigma_S^2 = \left(\frac{3}{11}\right)^2 \sigma_X^2 \\ &\cong \sigma_S^2 + \left(\frac{3}{11}\right)^2 10^{-1.5} \sigma_S^2 \\ &\cong 1.0024 \sigma_S^2\end{aligned}$$

This is important in relatively short-term time scales, typically less than 10 to 100 sec, depending on the magnitude of the Earth's ionosphere and interplanetary media-induced instabilities. Thus, for events that occur relatively quickly, such as small-scale vertical structures in the ionosphere, it is very important to maximize the S-band SNR. This is also true in terms of spectral density because layers in the ionosphere will induce multiple signals analogous to those described in the neutral atmosphere, and it is desirable to recover all the signals to recover the complete vertical profile of the ionosphere.

It is worth noting that this effect is very important in the removal of ionospheric effects, for instance, in the neutral atmospheric occultations. If, for example, the ionospheric effects are estimated and subtracted from the X-band neutral atmospheric data over time scales where this SNR effect is important, a tremendous increase in the short-term jitter of the X-band phase data will result by adding in phase noise from the lower SNR S-band signal. This effect must be considered before applying the ionospheric correction because in terms of ultimate measurement sensitivity, this cure may be worse than the problem.

For time scales greater than 10 to 100 sec, the S-3/11X-band phase is limited both by the equipment and any dispersive media between the Neptune system and the ground-based receivers. Media calibration is therefore very desirable, if available at a sufficient level of accuracy. In addition, in the case of neutral atmospheric occultations, the signal passes through the ionosphere on its way into and out of the neutral atmosphere, which underscores the importance of separating dispersive ionospheric phase shifts from neutral atmospheric effects when measuring the ionospheres of Neptune and Triton.

C. Rings

The SNR sets limits on both the minimum and maximum measurable opacities, as well as on the uncertainty associated with the measurements. The smallest measurable effect is limited by jitter in the signal amplitude associated with the SNR. The largest measurable opacity is limited to the point when the coherent signal can no longer be detected. In practice, the radial resolution of the diffraction correction process is typically reduced to increase the SNR and decrease the opacity uncertainties to more useful levels [3].

The SNR is also extremely important in determining the measurement sensitivity of the nearly forward-scattered signals that are combined when available with the coherent signal opacities to place limits on particle-size distributions and densities. The forward-scattered power is incoherent and occurs over a small range of frequencies shifted relative to the direct coherent signal due to a combination of the spacecraft and ring particle velocities. To detect this power, the scattered signal power spectral density must be higher than the receiver thermal noise spectral density. This is important at both S- and X-band frequencies because differential forward scattering power places constraints on the particle sizes. This effect was not expected to be detectable at Neptune, but was also not expected at Uranus and was indeed present although not fully understood [5].

One of the main objectives of these occultations is to measure the radially dependent microwave opacity. As previously mentioned, because the signal is coherent, the diffraction effects can largely be removed, which results in ~ 10 -to- 100 -m level resolution, depending on occultation geometry. There are numerous effects limiting the ultimately achievable resolution, including navigation errors, spacecraft antenna beamwidths, SNR, knowledge of the geometry, and the coherence time of the measurement system.

In the diffraction removal process, the rings are treated as an azimuthally symmetric diffraction grating depending only on radial distance from the planet. The phase length of a signal traveling from the spacecraft through a particular ringlet and then to the Earth is known so that the signal phase associated with this ringlet can be estimated as a function of time. This phase pattern can then be correlated against the received signal spectrum to detect the complex opacity of the ringlet. The opacity is complex because it potentially alters both the phase and amplitude of the signal passing through it, and because the ability to perform the matched filter technique depends on the time span over which the oscillator's phase is coherent. Eventually, over a sufficiently long time span, the oscillator's phase will gradually lose coherence, which causes the model and data to become uncorrelated and to set a maximum time span over which the phase matching process will work. In [8], an approximate relationship between coherence time and Allan deviation is used to derive a simple and general relationship between the oscillator's stability and the limitation that it places on the resolution of the recovered opacity profiles. A more rigorous and exact analysis for the case of a white-noise frequency oscillator has been given in [7].

For the reasons already given, where possible, it is desirable that the measurement-system phase-noise spectral density not degrade the SNR. Again, a case in point pertains to the data taken at Parkes during the Uranus encounter. Phase noise was discovered in the Parkes data set within a few hertz on either side of the carrier, with a much higher level than the continuous thermal phase noise. The additional phase noise reduced Parkes to a much less effective antenna and rendered it virtually useless from the point of arraying phase information. The source of the problem was subsequently isolated to be a noisy multiplier in the first local oscillator of the receiver. Since the amplitude of the LO was unaffected by this frequency multiplier, the amplitude spectrum of the downconverted signal in the Parkes data set had the full sensitivity of the G/T of the antenna and was, in fact, combined with the DSS 43 data, which provided a 1 to 1.5 dB improvement in the SNR.² This, of course, also points out the importance of the amplitude noise spectral density.

Systematic modulation of the signal created in the process of the reception can create apparent systematic patterns in the diffraction-corrected radial opacity profiles, a fact also discovered during processing of the Uranus data acquired at Parkes. The Parkes problem at the Uranus

encounter led to tighter specifications on spurious signals in the receiving systems.

The longer term amplitude stability is particularly important for detection of the presence of tenuous rings and for comparing measurements of the same ring on either side of the planet. The ability to associate a dip in the received signal level as a feature in the occulting media depends directly on the instrument's gain stability. This technique again demonstrates the importance of multiple receiving sites that allow determination of whether the effect is common in multiple data sets.

IX. Phase-Stability Characterization

A significant advance in the preparations for the Neptune encounter was the characterization of the phase-stability requirements in terms of statistical figures of merit used in the frequency and timing world. Using Allan variance and phase-noise spectral density (also referred to as spectral purity), the entire range of time scales of interest from 10 kHz on either side of the carrier to integration times of 1000 sec can be specified. The use of standardized figures of merit developed out of stability requirements late in the implementation efforts for the Uranus encounter and the preparations for the Galileo gravitational wave search experiment, which allowed stability error budgets to be created and clear design trade-off decisions to be made. Systematic testing of the performance of the various assemblies, subsystems, and systems against the required performance was also made possible.

The Allan variance was used to cover time scales of 1 sec and greater. The Allan variance provides no information on the absolute accuracy of the frequency of a signal, but does characterize its fluctuations over various time scales [9]. Since the radio occultation experiments primarily involve relative changes in phase and phase rate over a certain range of time scales, the Allan variance is an appropriate statistic. It also has the advantages of being simple to calculate and of converging for most types of noise encountered in these systems.

The spectral density of phase noise was used to characterize phase-stability performance over subsecond time scales. Specifically, performance was specified in terms of the single sideband phase noise, $S_{\phi}(f)$, over the range from 1 Hz to 10 kHz away from the carrier. The same level was used to specify the level of any spurious signals relative to the carrier signal level as well as the spectral density of

²L. Tyler, personal communication, Jet Propulsion Laboratory, Pasadena, California, 1987.

amplitude noise that was used to characterize the subsecond fluctuations in the signal amplitude. The long-term fluctuations of signal amplitude are relatively constrained as compared with signal-phase fluctuations. Therefore, a figure of merit that filters out the long-term fluctuations in determining performance over a range of time scales, such as the Allan variance, was not required, and a regular variance was used to specify amplitude performance.

The actual phase-stability requirements for the encounter were set with the idea that the stability of the received signal should be limited by the combination of the spacecraft USO, the propagation noise imposed by the intervening media, and the telecommunications link performance. The phase noise inherent in the downconversion process of the ground receiving system was to be essentially negligible as compared with the three sources of noise mentioned above. This goal was technically feasible partly because the USO was a 15-year-old crystal oscillator and better oscillators had since become available and partly because the distance to Neptune was so great that the received SNR was relatively low. "Essentially negligible" phase noise was chosen to mean 10 dB or more below the noise-power spectral density inherent in the link—a value chosen because a 10-percent increase in power increases the rms by a factor of only 5 percent and because it was committable as a guaranteed level of performance.

For stability characterizations using amplitude rather than power, such as the square root of the Allan variance or Allan deviation, the margin was chosen to be 3, or approximately the square root of 10. The USO stability³ and the corresponding system stability requirement, as characterized by the Allan deviation, is shown in Fig. 4. At 10 sec, performing a root sum square on the USO and system performance yields a combined Allan deviation of 1.044×10^{12} , or a degradation of 4.4 percent of the typical stability of the USO.

In the case of phase-noise spectral density, the signal at the output of the LNA was limited by a combination of the USO and the thermal phase noise resulting from the finite SNR. The estimated single-frequency performance and Voyager requirements are shown in Fig. 5. The maximum frequency of 10 kHz was chosen because the bandwidth of the receiver output was 20 kHz. The estimated stability of the received X-band signal is labeled VGR-N X-band 70 m, and it represents the combined limitations

of the X-band SNR and the Voyager USO. The overall X-band system performance was constrained to the levels labeled VGR-N X-band Requirement. The required S-band performance is less stringent than that of the X-band because of lower S-band SNR levels. Additionally, primarily due to cost, the hydrogen maser modifications were specified to meet the estimated Mars Observer requirements labeled as DSN Post-1989 X-band Requirement.

X. Coherent Arraying Requirements

In order to array two signals with near optimum combined SNRs, the relative phase difference between the two signals must be held within a small fraction of a cycle. This relative phase difference will drift because of the finite phase stability of the measuring systems as well as media effects along the different signal paths to each antenna on the Earth. The phase instability of the USO is common and, therefore, will not affect the relative phase difference. The desire is to hold the relative phase within this bound for as long as possible. In the extreme case, the ability to "blindly" array the signals received at two different antennas must be possible over the atmospheric occultation period. The need for this level of stability was driven by the possibility that the occulted S-band signal might be difficult, if not impossible, to detect upon reception at a single station, but the combined data would sufficiently increase the SNR to provide a reliable detection. The Allan variance is a relevant figure of merit for the stability required in this case.

One method of maintaining a small phase difference between received signals is to simply perform a linear phase extrapolation based on the measured phase difference in the past. One begins with the measured phase, at a time, τ , ago and runs a straight line through it and the measured phase at the present time. The error in the extrapolated phase estimate at a time, τ , in the future is related to the Allan variance as follows:

$$\sigma_{\phi \text{ ext}} = \sigma_y(\tau) \sqrt{2\tau} F_0$$

where $\sigma_{\phi \text{ ext}}$ is the rms of the extrapolated phase error in cycles, σ_y is the Allan deviation, and F_0 is the nominal downlink frequency. A similar argument can be made in the case of a simple linear phase interpolation where the phases at the beginning and end of an interval are known, but the signal phase is unknown within the interval corresponding to the situation case of an occultation. In this case the error in the phase estimate at the center of the interval can be shown to be:

³S. W. Asmar and P. M. Eshe, *Evaluation of the USO Performance—Final Report*, JPL IOM Voyager-RSST-90-121 (internal document), Jet Propulsion Laboratory, Pasadena, California, January 17, 1990.

$$\sigma_{\phi \text{ int}} = \frac{\sigma_y(\tau)\tau F_0}{\sqrt{2}}$$

where the interval length in this case is 2τ . As mentioned, the maximum interval of interest was approximately 2,000 sec, which covered the length of the Neptunian atmospheric occultation. Guaranteeing a phase error of less than 0.1 cycle at the center of this interval implies an rms phase error, $\sigma_{\phi \text{ int}}$, equal to 0.03 cycle. The equivalent σ_y at S-band would be about 2×10^{-14} . An analogous X-band number would be 5×10^{-15} . These are tight specifications relative to the performance and were treated as goals rather than as requirements.

Another useful aspect of these relationships is the determination of the duration over which the phase difference can be predicted with sufficient accuracy for arraying. It is particularly important to avoid adjusting the received phase over time scales where the signal phase contains relevant information. Measurement of small changes in the received signal phase caused by ringlets is a particular example. In that case, the coherent signal detection interval can be as long as the coherence time of the oscillators involved, which, in the case of Voyager, is approximately 100 sec at X-band.

XI. Differential S- Versus X-Band Phase Specifications

The actual S- versus X-band performance of the spacecraft radio frequency (RF) system is unknown since it was not measured with sufficient accuracy prior to launch. The RF system is presumed to be very stable due to its design and the relatively benign environment in which it resides.

An estimate of the interplanetary effects can be found in [10] where an extended span of dual frequency Viking data was used to estimate the range of phase instabilities caused by the interplanetary solar wind. The results indicate that the Allan deviation of the S-3/11X phase observable at the Voyager-Neptune encounter Sun-Earth-probe angle of ~ 127 deg ranges from about 3×10^{-14} to 2×10^{-13} .

Before choosing a specification, the needs of other spacecraft should be considered. The accuracy goal of the Galileo Faraday rotation experiment, which measures the difference between the orthogonal circular polarization components of the linearly polarized Galileo S-band downlink, is approximately the same value. The Faraday effect is caused by a combination of the magnetic field and electron density, requiring the S-3/11X phase measurement to characterize the electron density and allow recovery of the magnetic field from the differential S-band phase data.

The gravitational wave search experimenters also desire knowledge of the TEC in order to remove its effects from the signals, thereby improving instrument sensitivity. For the Galileo X-band uplink experiment, scheduled possibly as early as 1993, during the joint opposition with Mars Observer, measurements of the electron density using S-3/11X phase of the downlink can be used to reduce the plasma effects at least on the X-band downlink signal and possibly the uplink signal as well. The effect at opposition can potentially be as little as 2×10^{-14} at 1000 sec at S-band. The corresponding effect on the X-band downlink would be about 1.5×10^{-15} , which is a substantial portion of the 5×10^{-15} stability requirement on the entire X-band uplink and downlink system, so it is desirable to be able to accurately measure it and remove its effects at least from the downlink. An S-3/11X measurement accuracy of 6×10^{-15} would allow calibration of the X-band downlink to a level of 5×10^{-14} , which would make it negligible in the system stability error budget.

The actual Voyager specification on the S-3/11 X phase drift of the receiving equipment was 1 deg rms over 1000 sec,⁴ approximately equivalent to 1×10^{-15} . This was an old specification inherited from previous specifications, but never formally tested. In comparison with the interplanetary noise just described, it appears unnecessarily tight, but the open-loop receivers are designed for excellent common mode noise cancellation when the S-3/11X combination is formed. Therefore, the requirement was tentatively accepted because the system was expected to perform approximately at this level.

The total spectrum of system performance from 1 to 1000 sec is actually limited by the S-band SNR over the shorter time scales and by media noise over longer time scales, with a transition region in the 10 to 100 sec region, depending on the level of the media noise. This specification, reflecting both types of noise, is shown in Fig. 6.

XII. Antenna-Pointing Requirements

The basic desire for antenna pointing was to maximize the antenna gain over the period of interest without introducing significant phase and amplitude perturbations. Due to its unique location and mechanical design, each of the three antennas had a range of elevation angles over which the Neptune events would occur and a particular

⁴ *Deep Space Network and Flight Project Support Office Detailed Requirements for the Voyager Neptune/Interstellar Mission*, Project Document 618-858 (internal document), Jet Propulsion Laboratory, Pasadena, California, April 15, 1987.

maximum gain versus elevation angle curve of potential performance. The requirement for antenna-gain performance and, therefore, for pointing accuracy was that the antennas should point accurately enough so that their actual gain would be within 0.5 dB of their maximum potential gain as a function of elevation angle. A tighter requirement would have been desirable, but not practicable. In the interest of separating amplitude effects due to the antenna from those due to occulting media, a specification was also written that required the reconstructed antenna gain to be accurate to 0.1 dB.

A constraint that must immediately be considered is that the antenna pointing must be performed without aid of a signal feedback mechanism because of the dynamic nature of the received signal's amplitude and phase during occultations. The lack of feedback is significant because the DSN normally conically scans the antenna boresight around the direction of the spacecraft and uses the signal amplitude changes to keep the antenna on point once the signal is acquired.

The signal's optical path length through the antennas is significant. Path length changes due to the antenna were to be far less than the effects of interest. Phase shifts as small as 10 deg were apparent in the X-band data taken during the ring occultations at Uranus [5]. The phase shift is equivalent to a path length change of 1 mm, which resulted in a goal to understand any path-length changes in the antenna of 0.1 mm. The time scale of these ring effects is about 10 sec or less, which created a relatively tight specification of 3.3×10^{-14} over 10 sec at X-band. In reality, a requirement could have been set a factor of three higher and could still have kept within the total system stability specification of 3×10^{-13} over 10 sec. Relatively quick changes in path length are the real source of concern. A 0.1-mm change over 100 sec or more would never be apparent because of other sources of phase error.

An additional operational goal was the removal of the need for any special pointing calibrations. Normally, the subreflector position changes with time to keep the subreflector at the antenna focus as the pointing angles change. In the pointing mode chosen for the Uranus encounter, the subreflector position was fixed in all axes. Among other reasons, this was done to avoid any sudden path-length changes. While this strategy seemed to be safe and conservative, it resulted in a pointing offset relative to the standard spacecraft tracking mode, which made the standard pointing offset tables unusable and required special calibration tracks to generate the special offset tables. The multiple offset tables caused operational confusion. The subreflector positioning strategy ultimately used for

the Neptune occultation events allowed the subreflector to move along the Y-axis (effective elevation angle), which not only removed the need for special calibration tracks, but also introduced less path-length change than fixing the subreflector in all axes [11].

XIII. Follow-On Noise Temperature

Because the G/T of a deep-space receiving system is a very precious commodity, a goal was set that the elements in the receiving system following the LNA should degrade the front-end system temperature by no more than one percent. Given a front-end system temperature of approximately 20 K, this is equivalent to about 0.2 K in noise temperature. The primary sources of noise were identified as the receiver and the digitization process.

The receiver has a certain amount of inherent loss in transporting the signals over the long distances from the antenna front-end area to the control room. The receiver must also provide a sufficient gain from the levels at the LNA output to drive the analog-to-digital converters (ADC's). The additional requirement was set so that the noise of any amplifier stage in the receiver would be at least 23 dB below the amplified front-end noise at that stage of the receiver.

In the case of the signal digitization prior to recording, the resolution of the ADC used to digitize the down-converted signal can be determined by the requirement that the quantization noise be at least 23 dB below the front-end noise level at the receiver output. Taking the X-band signal-to-noise spectral density ratio to be 46 dB and the minimum final filter passband to be 20 kHz, or about 43 dB relative to 1 Hz, the total SNR of the down-converted and filtered signal would be $46 - 43 = 3$ dB.

To demonstrate the effect of the quantization noise, take an ADC with N bits of resolution resulting in 2^N levels equally spaced across the full-range voltage, FR , of the ADC. Next, assume that the quantization error is uniformly distributed from $-1/2$ lsb to $+1/2$ lsb, where an lsb is a least significant bit and equals FR divided by 2^N . This is a safe assumption, given that the front-end noise is well above that of the ADC quantization error, which is the goal. This assumption leads to an rms error of quantization noise of $\text{lsb}/\sqrt{12}$. In terms of its full range, the quantization rms noise level is then $FR/(2^N\sqrt{12}) = FR/(2^{N+1}\sqrt{3})$. If the input to the ADC were a sinusoid signal with amplitude, A_s , equal to half the full range of the ADC, the signal would just fit within the ADC's voltage range. The rms voltage of the sinusoid

would be $A_s/\sqrt{2}$ or $FR/2\sqrt{2}$. If the input sinusoid had a finite SNR equal to σ_S^2/σ_N^2 , the noise power of the input could also be expressed in terms of FR . For example, an SNR of 3 dB, the approximate X-band SNR in a 20 kHz bandwidth, results in $\sigma_N^2 = \sigma_S^2/2 = FR^2/16$. In reality, the input signal amplitude would never be set at the limits of the ADC because any additional noise would cause the input waveform to be clipped. To reduce the probability of clipping to an insignificant level, the input signal is attenuated by a factor of 16, which makes the input noise power equal to $FR^2/256$ or $FR^2 \times 2^{-8}$. The ratio of the input noise to the quantization noise becomes

$$\frac{FR^2 \times 2^{-8}}{FR^2/(3 \times 2^{2N+2})} = 3 \times 2^{2N-6}$$

Maintaining this ratio above 200 (23 dB) implies that N must be greater than 6. In fact, the ADCs of the existing system have 8-bit quantization, which results in an SNR increase of 0.03 percent, or 0.0014 dB, and which meets the "negligible" criterion.

XIV. Conclusion

The requirements for the radio science ground data system for the Voyager-Neptune encounter were developed based on experience with the system's performance at previous encounters, the unique geometry and configuration at Neptune, and the need to accommodate an array of three Pacific-basin ground stations. The specifications were derived from the Voyager Project's *Science Instrumentation Requirements Document*, as well as the DSN goal to accommodate, where appropriate, requirements for future missions. For the occultation experiments addressed here, the key requirement was acquisition of the spacecraft signal spectrum of interest in a phase-continuous manner with minimal degradation under dynamic conditions. This acquisition was accomplished by using open-loop receivers with excellent phase and frequency stability. The requirements discussed here emphasize the importance of system stability as defined in a clear set of specifications against which the system could be tested and the relation of stability to the ultimately derived quantities to characterize the Neptune system.

Acknowledgments

The authors thank North Ham for his support in the course of developing the radio science requirements, and Carole Hamilton for her review and comments on this article.

References

- [1] G. Fjeldbo, A. J. Kliore, and V. R. Eshleman, "The Neutral Atmosphere of Venus as Studied With the Mariner V Radio Occultation Experiments," *The Astronomical Journal*, vol. 76, no. 2, pp. 123-140, March 1971.
- [2] G. F. Lindal, J. R. Lyons, D. N. Sweetnam, V. R. Eshleman, D. P. Hinson, and G. L. Tyler, "The Atmosphere of Uranus: Results of Radio Occultation Measurements With Voyager 2," *Journal of Geophysical Research*, vol. 92, no. A13, pp. 14987-15001, December 30, 1987.
- [3] E. A. Marouf, G. L. Tyler, and P. A. Rosen, "Profiling Saturn's Rings by Radio Occultation," *Icarus*, vol. 68, pp. 120-166, 1986.
- [4] G. L. Tyler, E. A. Marouf, R. A. Simpson, H. A. Zebker, and V. R. Eshleman, "The Microwave Opacity of Saturn's Rings at Wavelengths of 3.6 and 13 cm From Voyager 1 Radio Occultation," *Icarus*, vol. 54, no. 2, pp. 160-188, May 1983.

- [5] D. L. Gresh, E. A. Marouf, G. L. Tyler, P. A. Rosen, and R. A. Simpson, "Voyager Radio Occultation by Uranus' Rings. I: Observational Results," *Icarus*, vol. 78, no. 1, pp. 131-168, 1989.
- [6] D. Neff, "Use of a 2.3 GHz Traveling-Wave Maser on the Usuda 64-Meter Antenna," *TDA Progress Report 42-89*, vol. January-March 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 34-40, May 15, 1987.
- [7] B. S. Haugstad, "Effects of the Inhomogeneous Background on Radiation Propagating through Turbulent Planetary Atmospheres," *Radio Science*, vol. 13, pp. 435-440, 1978.
- [8] E. R. Kursinski, "Application of High Stability Oscillators to Radio Science Experiments Using Deep Space Probes," *Proc. of the 22nd Annual Precise Time and Time Interval Planning Meeting*, Tysons Corner/Vienna, Virginia, December 4-6, 1990.
- [9] J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullen, W. L. Smith, R. L. Sydnor, R. F. C. Vessot, and G. M. R. Winkler, "Characterization of Frequency Stability," *IEEE Transactions on Instrumentation and Measurement*, vol. IM-20, no. 2, pp. 105-120, May 1971.
- [10] J. W. Armstrong, R. Woo, and F. B. Estabrook, "Interplanetary Phase Scintillation and the Search for Very Low Frequency Gravitational Radiation," *The Astrophysical Journal*, vol. 230, pp. 570-574, June 1, 1979.
- [11] S. D. Slobin and D. A. Bathker, "DSN 70-Meter Antenna X-Band Gain, Phase and Pointing Performance, With Particular Application for Voyager 2 at Neptune Encounter," *TDA Progress Report 42-95*, vol. July-September 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 237-245, November 15, 1988.
- [12] D. W. Brown, H. W. Cooper, J. W. Armstrong, and S. S. Kent, "Parkes-CDSCC Telemetry Array: Equipment Design," *TDA Progress Report 42-85*, vol. January-March 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 85-110, May 15, 1986.
- [13] D. L. Gray, S. E. Matousek, K. Francis, C. L. Potts, and R. J. Cesarone, "Voyager 2 Neptune Navigation Results," *AIAA/AAS Astrodynamics Conference*, paper AIAA 90-2876, Portland, Oregon, August 20-22, 1990.
- [14] S. D. Slobin and W. A. Imbriale, "DSS-43 Antenna Gain Analysis for Voyager Uranus Encounter: 8.45 GHz Radio Science Data Correction," *TDA Progress Report 42-90*, vol. April-June 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 127-135, August 15, 1987.

Table 1. Bandwidth allocation requirements

Radio source	Using LEU, kHz	Using LSU, kHz
Ring scattering	31.1	15.5
Atmosphere	21.8	8.4
Triton	29.5	13.9

Table 2. Signal-to-noise ratio comparison relative to DSS 43 64-m performance at Uranus encounter

	X-band, dB	S-band, dB
70-m extension	1.9	0.8
Antenna pointing ^a	2.0	0.2
Listen-only mode	—	1.0
Parkes [12]	-1.0	—
Usuda	—	0
Array at Neptune ^b	6.3	4.1

^a Due to DSS 43 antenna-pointing problems with the subreflector during Uranus encounter, degradation was -2 and 0.2 dB at X- and S-band, respectively [14].

^b Improvement in the SNR due to the Parkes/DSS 43 (70-m) and Usuda/ DSS 43 (70-m) arrays relative to DSS 43 alone at 64 m.

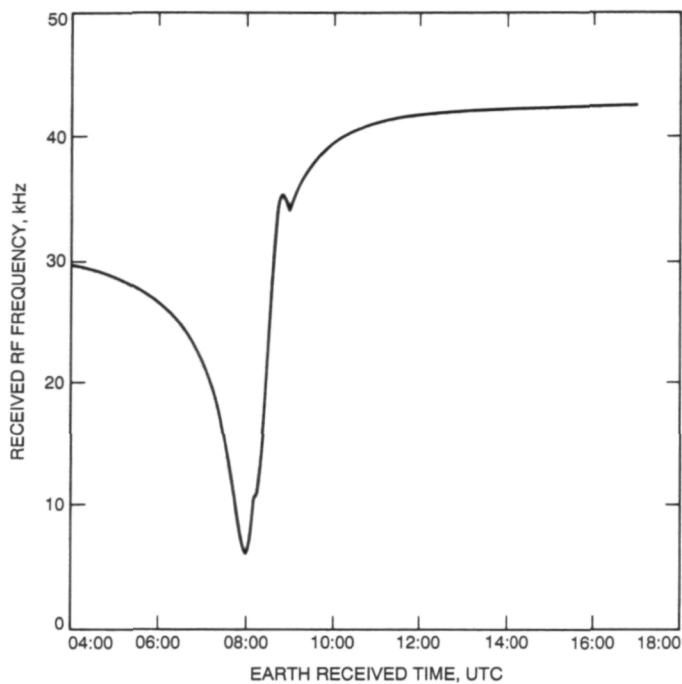


Fig. 1. DSS 43 X-band downlink frequency at 8419 MHz.

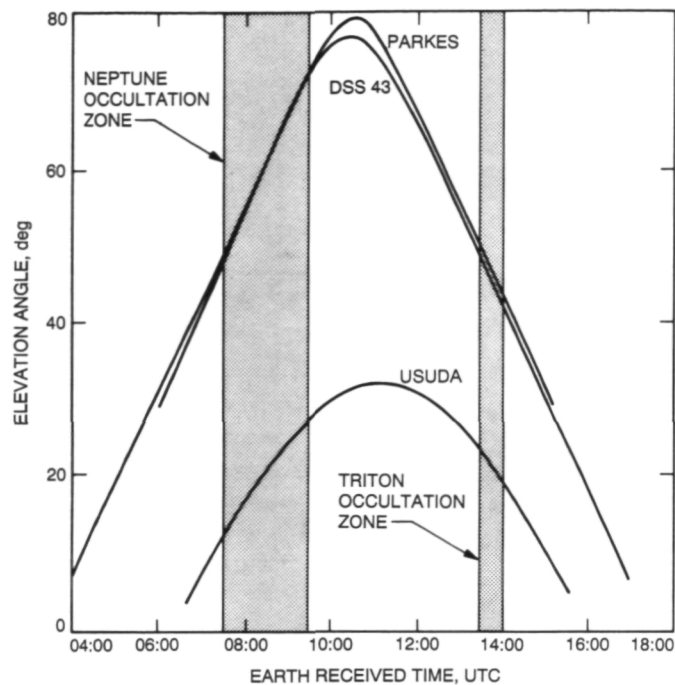


Fig. 3. DSS 43, Parkes, and Usuda elevation angles.

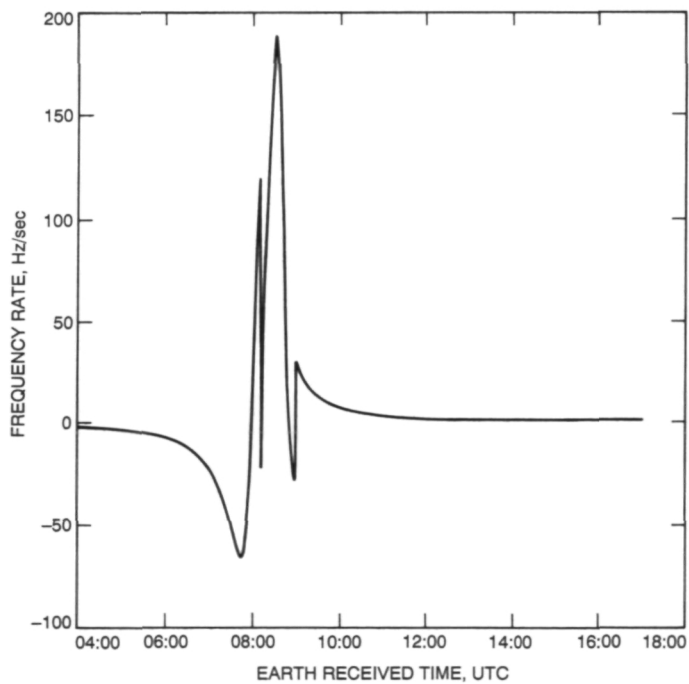


Fig. 2. DSS 43 X-band Doppler rates.

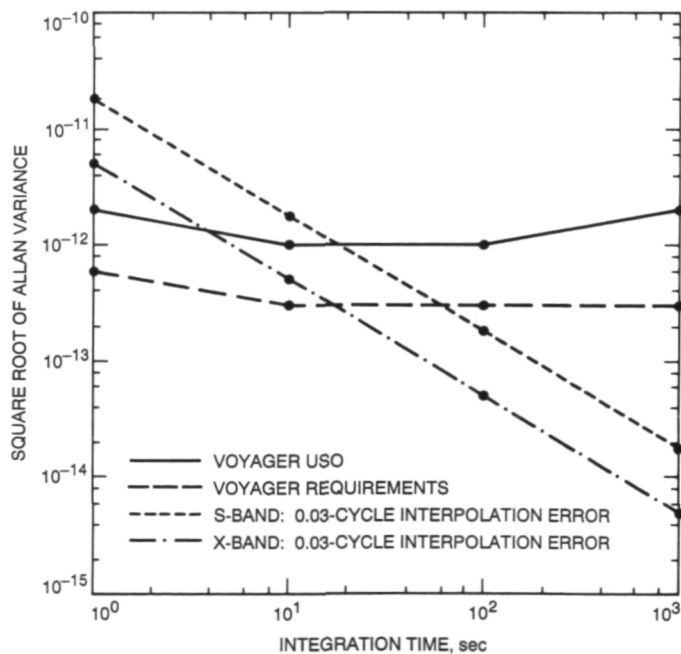


Fig. 4. Phase-stability performance.

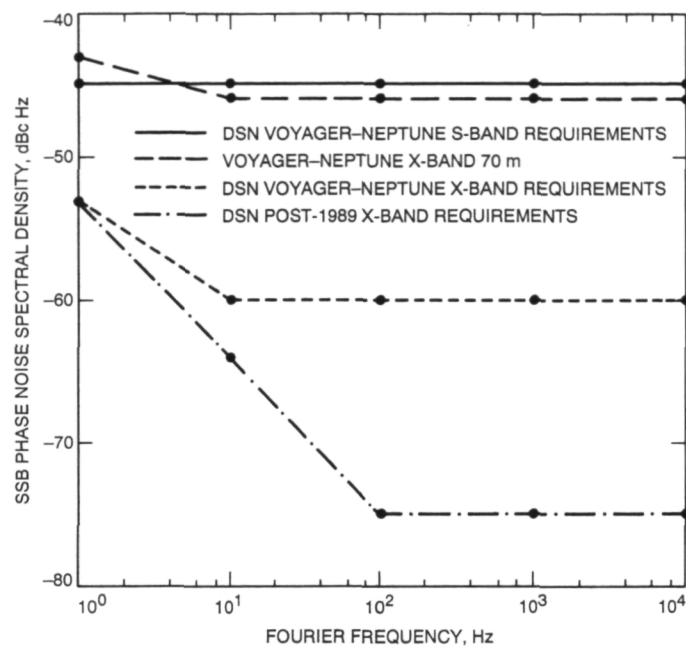


Fig. 5. Phase-noise performance.

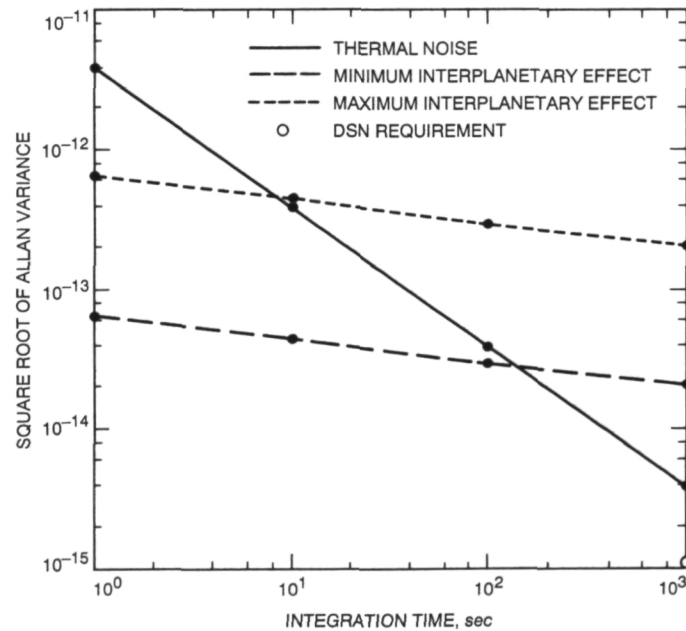


Fig. 6. Differential phase (S-3/11X) stability. This assumes a 30-dB-Hz SNR_0 and 1-Hz detection bandwidth. Interplanetary effects are derived from [10].

Gaussian Beam and Physical Optics Iteration Technique for Wideband Beam Waveguide Feed Design

W. Veruttipong, J. C. Chen, and D. A. Bathker
Ground Antennas and Facilities Engineering Section

The Gaussian beam technique has become increasingly popular for wideband beam waveguide (BWG) design. However, it is observed that the Gaussian solution is less accurate for smaller mirrors (approximately $< 30\lambda$ in diameter). Therefore, a high-performance wideband BWG design cannot be achieved by using the Gaussian beam technique alone. This article demonstrates a new design approach by iterating Gaussian beam and BWG parameters simultaneously at various frequencies to obtain a wideband BWG. The result is further improved by comparing it with physical optics results and repeating the iteration.

I. Introduction

Geometrical optics (GO) is a well-known technique used in the design of many beam waveguide (BWG) feed systems. A BWG feed system is composed of one or multiple feedhorns with a series of flat and curved mirrors arranged so that power can be propagated from the horn through the mirrors with minimum losses. Horns and equipment can thus be located in a large, stable enclosure at an accessible location. While GO is useful for designing high-frequency or electrically large mirrors (approximately $> 50\lambda$ in diameter with -20 dB edge taper), some BWGs may be operated at low frequency and have a mirror size of only about 20λ in diameter. Due to diffraction effects, the characteristics of a field propagated between small BWG mirrors (approximately $< 20\lambda$ in diameter) will be substantially different from the GO field. Therefore, the GO design is not suitable for a high-performance

wideband BWG antenna with small BWG mirrors. The Gaussian beam technique has become increasingly popular for wideband BWG design. The Gaussian beam mode is an approximate solution of a wave equation describing a beam of radiation that is unguided but effectively confined near an optical axis. The zero-order mode is normally used in the design. A major advantage of the Gaussian technique is the simplicity of the Gaussian formula, which is easy to implement with negligible computer time.

G. Goubau gave the first mathematical expression of Gaussian modes derived from the solution of Maxwell's equations described by a continuous spectrum of cylindrical waves [1]. T. S. Chu developed the Fresnel zone imaging principle of the Gaussian beam to design a pseudo-frequency-independent BWG feed [2]. S. Betsudan, T. Katagi, and S. Urasaki used a similar imaging tech-

nique to design large ground-based BWG antennas [3]. N. J. McEwan and P. F. Goldsmith developed a simple design procedure based on the Gaussian beam theory for illumination of reflector antennas where the reflector is electrically small or in the near-field of a feed [4].

A comparison of scatter fields calculated by zero-order Gaussian and physical optics (PO) solutions indicates that the Gaussian solution is less accurate for smaller mirrors (approximately $< 30\lambda$ in diameter). Therefore, a high-performance wideband BWG design may not be achieved by using the Gaussian beam technique alone. The purpose of this article is to demonstrate a new design approach by iterating Gaussian beam and BWG parameters to obtain a wideband BWG feed. The result is further improved by comparing the Gaussian beam results with the PO results and repeating the iteration. Details will be described next.

II. Design Geometry

The basic goal is to design a BWG feed system with good performance from 2 to 32 GHz by utilizing mirror sizes of 20λ in diameter with edge taper -23 dB at 2.295 GHz (other frequencies may have different edge tapers). It is also required that all feedhorns be placed in a basement room below the azimuth wheel and track. The feed system must provide a simultaneous operation capability and a fast response feed selection system. The geometry of the BWG antenna is shown in Fig. 1, where M_2 , M_3 , and M_5 are curved mirrors and M_1 , M_4 , and M_6 are flat plates. If M_6 is a dichroic plate, an additional reflector and horn, beyond M_6 , can provide dual frequency simultaneous operation. All the flat plates are assumed to be sufficiently large and are excluded from the design. The BWG in Fig. 1 can be reduced to a horn and three curved mirrors (Fig. 2) with curved mirrors replaced by thin lenses. Let f_i be the focal length and let D_i be the diameter of lens M_i with edge taper $-T_i$ dB ($i = 2, 3$, and 5). The horn and mirrors are separated by distances L_i ($i = 0, 1, 2, 3$), with L_H representing the length of the horn. D_H and D_{SUB} are diameters of the horn and the subreflector, respectively. R and D are the wavefront radius of curvature and the -18 -dB-beam diameter of the Gaussian beam at the subreflector, respectively. The -18 -dB-beam diameter is defined as the diameter at which the field amplitude has fallen 18 dB from its maximum value.

III. Design Procedure

Different frequencies usually have different values of D . The design goal is to have R and D be constant over the

design frequency range with acceptable spillover losses at all mirrors and $D = D_{SUB}$. Three frequencies: 2.295 GHz (S-band), 8.45 GHz (X-band), and 32.0 GHz (Ka-band) are used in the design. Curved mirrors are arranged so that the mirror system has a low cross-polarization. For a different design frequency range, the S-, X-, and Ka-bands can be replaced by low, middle, and high frequencies of the band. Input parameters are operating frequencies, D_{SUB} , D_2 , D_3 , D_5 , and maximum allowable spillover loss (or dB edge taper) at each mirror. The relationship between D_H and R_H is known. The rest of the parameters in Fig. 2 are unknown and are to be determined during the iteration process. The desired ranges of some of the input and output parameters need to be established. The initial values to start the iteration can be obtained by the GO design that roughly fits to antenna structures or other requirements.

The design procedure can be described as follows:

Step 1. The radius of curvature R and beam diameter D at the subreflector are calculated starting from the horn and proceeding through mirrors M_5 , M_3 , and M_2 to the subreflector by using the zero-order Gaussian mode. Details are shown in the Appendix. Let R_s , R_x , and R_{ka} (D_s , D_x , and D_{ka}) be the radii of curvature (-18 -dB-beam diameters) at the subreflector calculated at S-, X-, and Ka-bands, respectively. The unknown parameters are iterated so that $R_s = R_x = R_{ka}$ and $D_s = D_x = D_{ka}$. It is quite easy to have $R_x = R_{ka}$ (as well as $D_x = D_{ka}$). However, in many cases (due to structure constraints, size of mirrors, etc.), the iteration cannot converge to the condition $R_s = R_x = R_{ka}$. Instead, R_s is usually greater than R_x and R_{ka} . Therefore, one might have to accept $R_s > R_x = R_{ka}$.

Step 2. The (R_s, R_x, R_{ka}) and (D_s, D_x, D_{ka}) are recalculated by PO, a more accurate technique, with BWG parameters obtained from Step 1. Recall that the Gaussian solution predicts that $R_s = R_x = R_{ka}$ in Step 1, while PO results show that $R_s > R_x > R_{ka}$. The beam diameters from PO calculations are slightly smaller than the Gaussian results at all frequencies (but the trend may not be consistent for other cases). It is noted that the differences of R calculated from PO and Gaussian software are larger at electrically smaller mirrors.

Step 3. In order to offset the discrepancy between Gaussian and PO results, as indicated in Step 2, Step 1 is repeated and the unknown parameters are iterated so that $R_s < R_x < R_{ka}$, which are approximately the same amounts as indicated in Step 2 but in the opposite sense ("larger" in Step 2 results in "smaller" in Step 3). A nu-

merical example will be provided shortly. A similar adjustment procedure is also applied for the beam diameters D_s , D_x , and D_{ka} .

Steps 2 and 3 can be repeated until an acceptable result is achieved. For simplicity, only radii of curvature at X- and Ka-bands are considered as examples here. In Step 1, after millions of iterations, one obtains $R_x = R_{ka} = 476$ in. The calculation in Step 2 by PO software gives $R_x = 488$ in. and $R_{ka} = 478$ in. with $\Delta R_x = 488 - 476 = 12$ in. and $\Delta R_{ka} = 478 - 476 = 2$ in. In Step 3, the goal is to iterate the unknown parameter so that $R_x = 476 - 12 = 464$ in. and $R_{ka} = 476 - 2 = 474$ in. When Step 2 is repeated with parameters recently obtained from Step 3, the results are $R_x = 477$ in. and $R_{ka} = 476$ in. The radius of curvature $R = 476.5$ in. is chosen for a dual-shaped reflector synthesis.

Table 1 shows a numerical comparison between Gaussian and PO techniques of the BWG configuration shown in Fig. 3. If the PO software is not available, one could use data from Table 1, provided that the new BWG configuration closely resembles the one in Fig. 3. It is noted that discrepancies between PO and Gaussian results are larger for electrically smaller diameters. The discrepancies are less sensitive to the distance between mirrors as long as they are in the Fresnel zone. In a design with the minimum mirror diameter $> 50\lambda$, reasonably good performance can be achieved by implementing only Step 1.

IV. Conclusion

The result from this design technique is shown in Fig. 3, with $f_2 = f_3 = 2580$ in. and $f_5 = 220$ in. The mirror diameters are $D_2 = D_3 = 105$ in., $D_5 = 131.5$ in., with their average edge tapers shown in Table 1. It is noted that the design result shown in Fig. 3 is close to the optimum performance. Some small performance sacrifices are made for cost, structure retrofit, maintenance, and accessibility. Also, some S-band performance is sacrificed in order to achieve very good performances at X- and Ka-bands. Spillover loss at each mirror is listed in Table 2. The spillover loss is calculated by integrating a scattered field calculated by the PO software, with the assumption that there is no tube effect. It is observed from Table 2 that energy is more confined well inside the BWG at higher frequencies since the spillover loss is lower at higher frequencies. Higher gain horns are needed for higher frequencies. Aperture diameters of S-, X-, and Ka-band horns are 4.57λ , 10.02λ , and 15.44λ , respectively. One operating mode for this antenna is simultaneous S-/X-band, with the dichroic surface M_6 reflecting S-band and passing X-band. In another operation mode, M_6 is flipped out of the beam path, allowing Ka-band (32.0 GHz; as well as X-band) to propagate to M_7 . Given that M_7 is a dichroic surface, simultaneous X-/Ka-band operation is achieved. By simply rotating M_5 , extra frequency bands can be used. In 1993, the NASA Deep Space Network (DSN) will begin construction of three 34-m BWG antennas based on this high-performance design concept.

References

- [1] G. Goubau and F. Schwering, "On the Guide Propagation of Electromagnetic Wave Beams," *IRE Trans. Antennas Propagat.*, vol. AP-9, no. 5, pp. 248-256, May 1961.
- [2] T. S. Chu, "An Imaging Beam Waveguide Feed," *IEEE Trans. Antennas Propagat.*, vol. AP-31, no. 4, pp. 614-619, July 1983.
- [3] S. Betsudan, T. Katagi, and S. Urasaki, "Design Method of Four Reflector-Type Beam Waveguide Feeds," *Japanese Electronics and Communications Society Journal*, vol. J67-B, no. 6, pp. 622-629, June 1984.
- [4] N. J. McEwan and P. F. Goldsmith, "Gaussian Beam Techniques for Illuminating Reflector Antenna," *IEEE Trans. Antennas Propagat.*, vol. 37, no. 3, pp. 297-304, March 1989.
- [5] T. S. Chu, "Geometrical Representation of Gaussian Beam Propagation," *Bell Syst. Tech. Journal*, vol. 45, no. 2, pp. 287-299, February 1966.

Table 1. Radius of curvatures and beamwidths of fields at the subreflector and average edge tapers of M_2 , M_3 , and M_5 , as shown in Fig. 3

Frequency band	R , in.		D , in.		Average edge taper, dB		
	Gaussian	PO	Gaussian	PO	M_2	M_3	M_5
S	462.2	515.0	138.5	137.9	-22	-26	-21
X	464.1	477.0	136.9	134.5	-29	-28	-32
Ka	473.7	476.0	133.7	133.0	-30	-28	-33

Table 2. Spillover loss of each mirror at 2.295, 8.45, and 32.0 GHz

Mirrors	Spillover losses, dB		
	2.295 GHz	8.45 GHz	32.0 GHz
M_1	0.004	0.006	0.006
M_2	0.051	0.015	0.014
M_3	0.018	0.017	0.017
M_4	0.017	0.005	0.004
M_5	0.075	0.005	0.005
M_6	0.025	0.004	-
M_7	-	0.003	0.004
M_8	-	-	0.003
Total	0.190	0.055	0.053

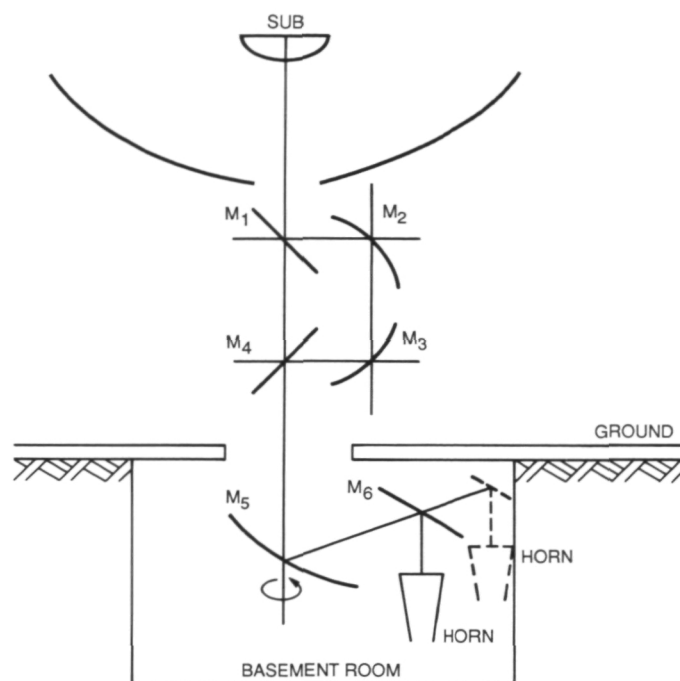


Fig. 1. Beam waveguide design configuration.

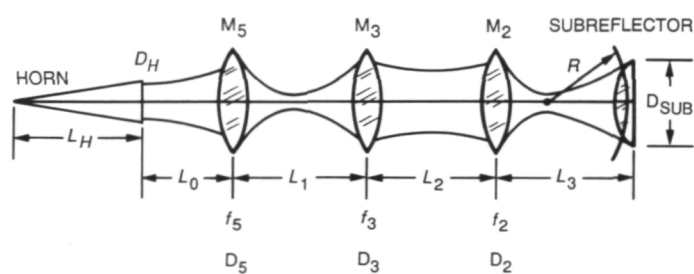


Fig. 2. Beam waveguide design parameters.

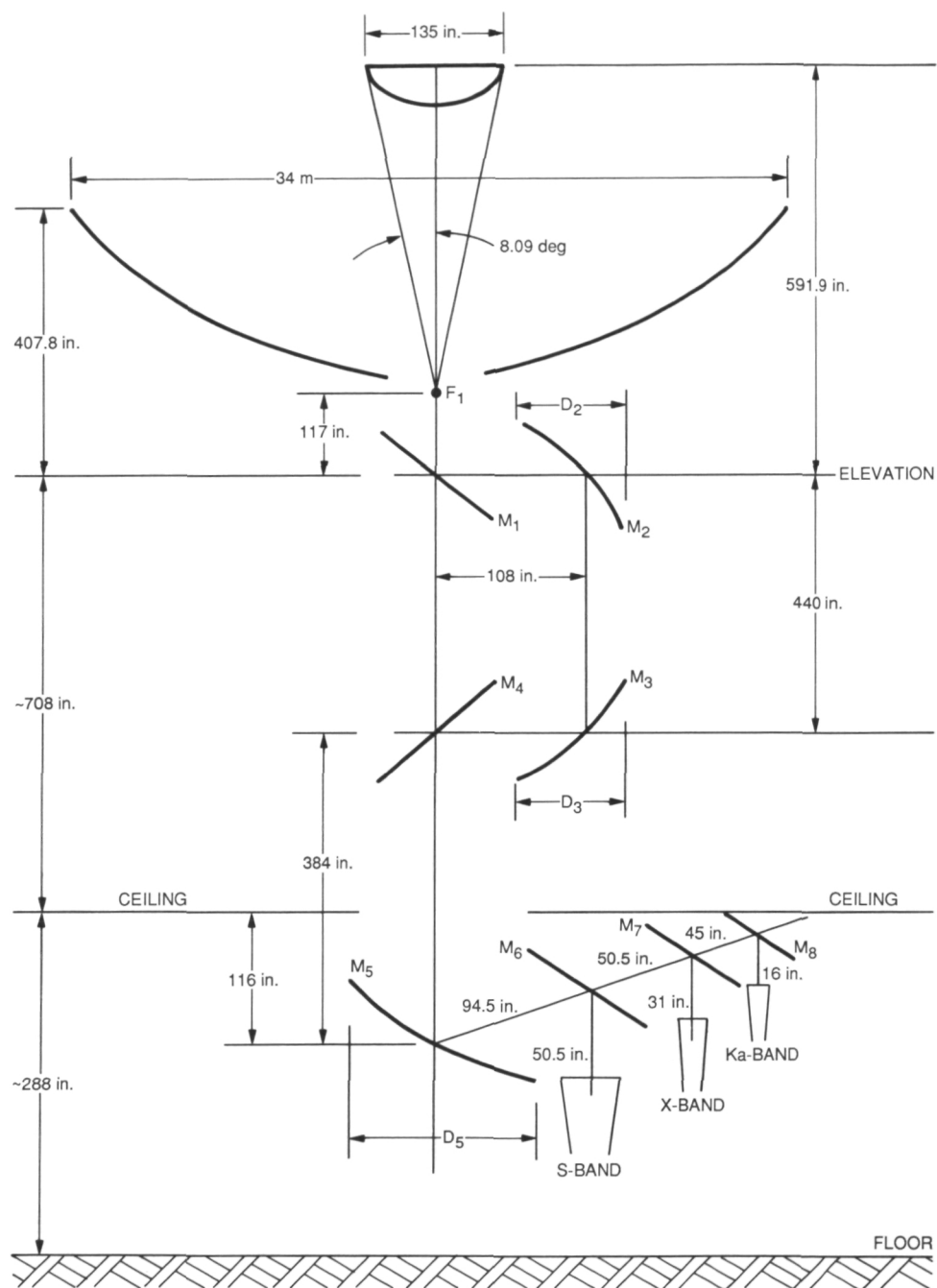


Fig. 3. Detailed dimensions of the beam waveguide configuration.

Appendix

The purpose of this appendix is to show zero-order Gaussian mode expressions used to compute fields from a horn through a series of mirrors and ending at a subreflector.

A. From a Horn to a Mirror (a Thin Lens)

The beam radius and phase front radius of curvature at the aperture of a corrugated horn are

$$\omega_0 = 0.32 D_H \quad (\text{A-1})$$

and

$$r_0 = \sqrt{L_H^2 + \frac{D_H^2}{4}} \quad (\text{A-2})$$

respectively [2]. The beam radius is defined as the radius at which the field amplitude has fallen 1/e of its maximum values, where D_H and L_H are the diameter and length of the corrugated horn, as shown in Fig. A-1. The radius of curvature on the lefthand side of the thin lens M_1 is [5]:

$$r'_1 = \frac{L_0}{\left[1 - \frac{1 + L_0/r_0}{\left(1 + L_0/r_0 \right)^2 + \left(\lambda L_0 / \pi \omega_0^2 \right)^2} \right]} \quad (\text{A-3})$$

The beam radius at M_1 (which is the same on both sides of the thin lens) is [5]:

$$\omega_1 = \omega_0 \sqrt{\left(1 + L_0/r_0 \right)^2 + \left(\lambda L_0 / \pi \omega_0^2 \right)^2} \quad (\text{A-4})$$

where L_0 is the distance from the horn aperture to the center of the lens (or mirror), and λ is a wavelength of an operating frequency.

The diameter and spillover loss of the lens M_1 with $-T_1$ -dB edge taper are each [3]

$$D_1 = 0.6786 \omega \sqrt{T_1} \quad (\text{A-5})$$

and

$$P_1 = 10 \log_{10}(1 - e^{-0.2303 T_1}) \quad (\text{A-6})$$

Spillover loss in Eq. (A-6) is reasonably accurate for design purposes. After the design is completed, more accurate predictions of the spillover loss will be obtained by direct integration of a scattered field computed by PO and spherical wave expansion software. The radius of curvature on the right-hand side of the lens M_1 can be obtained from a thin lens relation

$$\frac{1}{r_1} = \frac{1}{f_1} - \frac{1}{r'_1} \quad (\text{A-7})$$

where f_1 is the focal length of the lens M_1 , and r_1 (as well as r'_1) is defined to be positive when its phase front is convex toward the lens surface.

B. Between Mirrors

Similar to Section A, the radius of curvature and beamwidth at the left-hand side of M_2 are

$$r'_2 = \frac{L_1}{\left[1 - \frac{1 - L_1/r_1}{\left(1 - L_1/r_1 \right)^2 + \left(\lambda L_1 / \pi \omega_1^2 \right)^2} \right]} \quad (\text{A-8})$$

and

$$\omega_2 = \omega_1 \sqrt{\left(1 - L_1/r_1 \right)^2 + \left(\lambda L_1 / \pi \omega_1^2 \right)^2} \quad (\text{A-9})$$

where L_1 is the distance between the two lenses. ω_1 and r_1 are given in Eqs. (A-4) and (A-7), respectively. The differences between signs in Eqs. (A-3) and (A-8), and also in Eqs. (A-4) and (A-9) are due to a definition that a radius of curvature of a Gaussian beam is negative when it is concave toward the direction of propagation (+Z), and r_1 is concave toward +Z as in Fig. A-2. The value of r_1 obtained from Eq. (A-7) can be directly substituted into Eqs. (A-8) or (A-9) without changing any of the signs. The diameter and spillover loss of the lens M_2 for $-T_2$ -dB taper can be obtained from Eqs. (A-5) and (A-6), with T_1 replaced by T_2 . It is noted that M_2 can be a subreflector. The calculations in Section B are repeated for all the rest of the mirrors and the subreflector.

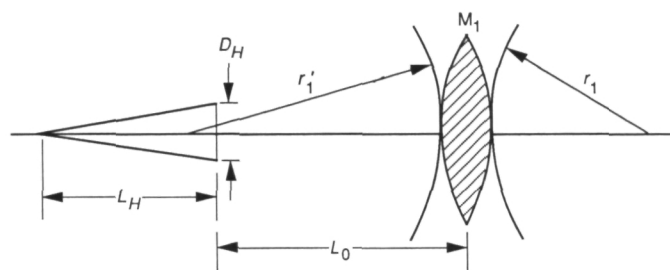


Fig. A-1. The geometry of a circular aperture corrugated horn illuminating a thin lens.

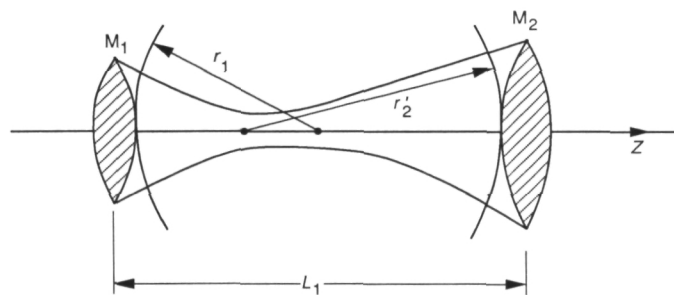


Fig. A-2. The geometry of a Gaussian beam between two lenses.

1994025122

N94-29625

TDA Progress Report 42-105

May 15, 1991

442620

Design and Analysis of a Low-Loss Linear Analog Phase Modulator for Deep Space Spacecraft X-Band Transponder Applications

N. R. Mysoor and R. O. Mueller
Spacecraft Telecommunications Equipment Section

This article summarizes the design concepts, analyses, and development of an X-band (8145 MHz) transponder low-loss linear phase modulator for deep space spacecraft applications. A single-section breadboard circulator-coupled reflection phase modulator has been analyzed, fabricated, and evaluated. A linear phase deviation of 92 deg with a linearity tolerance of ± 8 percent was measured for this modulator from 8257 MHz to 8634 MHz over the temperature range -20 deg C to $+75$ deg C. The measured insertion loss and the static delay variation with temperature were 2 ± 0.3 dB and 0.16 psec/deg C, respectively. Based on this design, cascaded sections have been modeled, and simulations were performed to provide an X-band deep space transponder (DST) phase modulator with ± 2.5 radians (± 143 deg) of peak phase deviation to accommodate downlink signal modulation with composite telemetry data and ranging, with a deviation linearity tolerance of ± 8 percent and insertion loss of less than 10 ± 0.5 dB. A two-section phase modulator using constant gamma hyperabrupt varactors and an efficient modulator driver circuit was breadboarded. The measured results satisfy the DST phase-modulator requirements and show excellent agreement with the predicted results.

I. Introduction

A circulator-coupled reflection phase modulator has been analyzed and investigated to provide the capability to directly modulate an X-band (8415 MHz) downlink carrier for the next generation of spaceborne communications systems. The phase modulator must be capable of large linear phase deviation, low loss, and wideband operation

with good thermal stability. In addition, the phase modulator and its driver circuit must be compact and consume low dc power. The design is to provide ± 2.5 radians (± 143 degrees) of peak phase deviation to accommodate downlink modulation of telemetry and ranging signals. The tolerance on the phase deviation linearity is ± 8 percent. The insertion loss should be less than 10 dB, and its variation with phase shift should be

within ± 0.5 dB. The phase delay variation specifications over the transponder hardware qualification environment, -20 deg C to $+75$ deg C, are less than 32 psec/deg C for the transponder, and less than 0.5 psec/deg C for the phase modulator. Such stringent specifications make the hardware implementation rather difficult. This investigation, which updates research described in [1], considers the reflection-type phase shifter for the implementation of the hardware. The results extrapolated from analyses and measured performance for a single-section phase modulator with high phase-resolution capability are presented in this article. Theoretical analyses of the modulator are presented in Section II, the breadboard modulator configuration and test data are presented in Section III, and the conclusions are presented in Section IV.

II. Phase Modulator Analyses

This investigation considers the circulator-reflection-type phase shifter. A single-stage phase shifter is shown in Fig. 1. Theoretical analyses of the single-stage and multi-stage phase-modulator circuits are presented in the following subsections. Both abrupt-junction and hyperabrupt-junction varactors will be considered.

A. Analysis of a Single-Stage, Reflection Phase Modulator

The varactor diode is well suited for the phase modulator [2,3] application because it can provide rapid phase change with the applied voltage. The circuit model for a packaged diode terminating a transmission line of characteristic impedance, Z_0 , for a reflective phase shifter is shown in Fig. 2. The junction capacitance of varactor tuning diodes is modeled as

$$\frac{C_j(V)}{C_0} = \left(1 + \frac{|V|}{\Phi}\right) - \Gamma$$

where $C_j(V)$ is the junction capacitance at reverse bias voltage, V . The absolute value of the applied reverse bias voltage is $|V|$. Additionally, C_0 is the junction capacitance at $V = 0$, and Φ is the built-in potential, which is equal to 0.8 V for silicon and 1.3 V for gallium arsenide. The diode capacity variation parameter, Γ , is the slope of the capacitance-voltage ($C - V$) curve when plotted on log-log scaled paper. This slope (Γ) is a function of the bias voltage and junction temperature. In the operating bias range of the diode, Γ can be treated as a constant.

1. Silicon Abrupt Junction Varactors. The diode capacity parameter Γ is equal to about 0.5 for practical abrupt-junction silicon diodes. For this analysis, an abrupt-junction silicon diode of capacitance C_j equal to 0.6 pF at -4 V bias is considered. Shown in the circuit model [1] of Fig. 2 are the diode junction capacitance, diode leakage resistance (R_s), package inductance (L_p), package capacitance (C_p), lead inductance (L_l), parallel resistance (R_p), and 10-ohm transmission line. To obtain a large linear phase shift, it is necessary to use a low-impedance transmission line at the diode terminal. The 10-ohm line was selected because it can be realized on a 0.254-mm-thick alumina substrate. Furthermore, its line width is not overly wide as compared with line length and the diameter of the device package. The line impedances and lengths are optimized to provide linear phase-shift variation with bias voltage. The low parasitic packages used in this investigation are Alpha Associates 304-001 and Microwave Associates MA-96 packages. The typical values of package parasitics for these packages are: $L_p = 0.15$ nH and $C_p = 0.17$ pF. The contact lead inductance $L_l = 0.05$ nH. The device series resistance R_s is a function of Q . The following expression for Q as a function of the bias voltage, V , was obtained by curve fitting the manufacturer's (Alpha Associates') data at 50 MHz:

$$Q(V) = 1700 + 875|V|$$

The diode series resistance is then given by:

$$R_s(V) = [2\pi \times (50 \times 10^6)QC_j(V)]^{-1}$$

The parallel resistance R_p at the diode plane is intentionally added to maintain the insertion loss constant. The calculated phase shift and insertion loss variations with the bias voltage for this model are shown in Fig. 3. The phase shift is linear over the 4.5 ± 2.5 V bias range. The linear phase shift over this range is ± 50 degrees. The maximum deviation from linearity is within ± 8 percent. The insertion loss variation over this voltage range is 1.8 ± 0.15 dB. This insertion loss does not include the losses due to the circulator and transmission line matching sections. The insertion loss of a phase modulator must be kept constant with phase shift. The variation of the insertion loss with phase shift causes amplitude modulation (AM) of the RF signal. This AM may be converted to undesirable phase modulation (PM) by subsequent nonlinear operation.

2. Gallium Arsenide Constant Gamma Hyperabrupt Varactors. The gamma of typical hyperabrupt varactors varies widely with applied voltage, which makes

these varactors unsuitable for linear phase-modulator applications. Recent processing developments have enabled the construction of varactors with gammas that remain constant over a limited voltage range. The varactors of constant Γ equal to 1.25 and 1.5 have been analyzed. Both were simulated by using the following package (MA-96) parasitics: $C_{-4V} = 0.6$ pF, $C_p = 0.17$ pF, and $L_p = 0.15$ nH. The expression for Q that was used in the simulation was obtained by curve fitting the manufacturer's (Microwave Associates') data at 50 MHz.

$$Q(V) = 1538 - 115.6|V| + 118.2|V|^2 + 21.96|V|^3 - 1.683|V|^4$$

The optimum characteristic impedance of the transmission line connected to the varactor is 30 ohms; this value provides flat insertion loss and large linear phase shift over the designed bias range. The simulation results for single-section hyperabrupt varactors are shown in Fig. 3. Linear operation is obtained at 4.5 ± 2.5 V, which results in a phase shift of ± 75 deg (< 8 percent linearity deviation) and an insertion loss of 1.3 ± 0.05 dB for a gamma of 1.5. The results also show that a gamma of 1.5 yields a loss that is 0.2 dB lower than a gamma of 1.25, although both have the same linearity and AM characteristics. The higher gamma value provides a more linear tuning characteristic, which results in a more sensitive phase modulator: 30 deg/V as compared with 20 deg/V for a gamma of 0.47.

B. Analysis of Multistage Reflection Phase Modulator

Comparisons of ± 143 -deg multisection phase shifters for various diode gamma values are given in Fig. 4 and Table 1. The results do not include the losses due to circulators and isolators and show that only two sections utilizing constant gamma hyperabrupt varactors have to be cascaded, as compared with three sections for the abrupt junction case. The predicted phase sensitivity for these cases is equal to 60 deg/V. The lowest insertion loss of 2.55 ± 0.1 dB is predicted for the two-section varactor modulator with a constant gamma of 1.5.

III. Phase Modulator Implementation and Performance

The circuit configuration and measured performances of the breadboard single-stage and two-section circulator reflection phase modulators are given in subsections III.A and III.B. The performance of the modulator driver circuit is presented in subsection III.C.

A. Single-Stage Circulator Reflection Phase Modulator

1. Circuit Configuration. The circuit diagram and a photograph of the phase modulator (with abrupt junction varactors) are shown in Figs. 1 and 5, respectively. The reflection phase shifter configuration [2,3] makes use of a 50-ohm circulator to provide matched input and output terminals for the phase-shifting diode circuit in the middle path. The circulator used in this investigation is a broadband 8.4 to 12 GHz circulator (Western Microwave, Inc., 13CX-481, Serial no. 10). The two-way insertion loss of the circulator is 0.6 ± 0.1 dB over the operating bandwidth at the nominal temperature of 23 deg C. The circuit (Fig. 1) consists of a packaged diode at the end of a 10-ohm line and two quarter-wave matching sections (33.44-ohm and 14.95-ohm sections) to transform 10 ohms at the diode terminals to 50 ohms at the circulator port. The 10-ohm microstrip circuit is etched on a 0.254-mm-thick-Roger 6010.5 substrate of dielectric constant $K = 10.5$. The width of the 10.0-ohm line is 2.464 mm, which is slightly (20 percent) larger than the diode package diameter of 2.03 mm. Better than 75 percent size matching between the diode and the line width is necessary in hybrid circuits to reduce insertion loss. The diode's anode is soldered to the ground as shown in Fig. 5. The circuit model for the diode, diode package parasitics, and connecting lead inductance are illustrated in Fig. 2. The junction capacitance of the selected abrupt-junction silicon diode (Alpha Industries, DVH 6733-02 in 168-001 package) is approximately equal to 0.6 pF at -4 V bias, 0.45 pF at -8 V bias, and 1.392 pF at 0 V bias. This results in a diode capacitance ratio of 3:1 from 0 V to -8 V bias range. Such large capacitance variation with bias is necessary to obtain large phase deviation. The series resistance of the diode is 3.6 ohms. The phase deviation characteristics of the circuit are also influenced by the values of package parasitics, interconnection lead inductance, and the length of the 10.0-ohm line. The typical values of the diode package parasitics supplied by the manufacturer are: $L_p = 0.5$ nH and $C_p = 0.18$ pF. The inductance of the interconnection lead (L_l) between the diode package and the 10.0-ohm line is equal to 0.05 nH. It is difficult to accurately model the package parasitics. The lengths of the interconnection lead (L_l) and the 10.0-ohm line can be adjusted to obtain linear phase deviation.

2. Performance of Single-Stage Phase Modulator at 8415 MHz. The measured phase deviation versus bias characteristics for the single-section phase modulator (Fig. 5) with an abrupt junction varactor is illustrated in Fig. 6. The nominal values of frequency, bias, voltage, and temperature in these measurements are 8415 MHz, $+4.5$ V and $+23$ deg C, respectively. The measurement band-

width ranged from 8257 MHz to 8634 MHz. The phase angle measured at 0 V bias, and the nominal frequency 8415 MHz was used as the reference angle. The measured linear phase deviation for the voltage swing ± 3 V above the nominal bias of 4.5 V was ± 46 deg with a linearity better than ± 8 percent of a best-fitted straight line. The phase-modulator circuit was subjected to temperature tests over the hardware qualification temperature range from -20 deg C to $+75$ deg C. As shown in Fig. 7, the overall variation of the static phase with temperature is about 45 deg with 0.5 deg/deg C slope. The calculated static phase delay at 8415 MHz is 0.16 psec/deg C. As seen from Fig. 7, the change in the static phase with temperature is not symmetrical about its value at $+23$ deg C. The reason is that the static phase variation of the circulator with temperature is nonlinear. The measured value of the circulator's static phase shift was 2.2 deg for a change in temperature from $+23$ deg C to $+75$ deg C, and was equal to -17.3 deg from $+23$ deg C to -20 deg C. But the temperature-induced phase shift for the diode circuit has a linear slope equal to 0.26 deg/deg C. Measured insertion loss as a function of varactor bias and temperature is shown in Fig. 7. The RF input power level was equal to 5 dBm at 8415 MHz. The maximum variation of the insertion loss with the circulator was found to be 2 ± 0.3 dB over the bias levels (4.5 ± 3 V), temperature range (-20 deg C to $+75$ deg C), and the RF range (8257 MHz to 8634 MHz). This includes the circulator's two-way insertion loss of 0.6 dB. The diode phase shift and linearity can be calculated if all the circuit component values are known fairly accurately. The values of package parasitics and the diode's junction dynamics are not well known, which results in a discrepancy between the measured and predicted results.

B. Two-Stage Circulator Reflection Phase Modulator

A two-stage phase modulator with a 1.5-gamma hyper-abrupt varactor was breadboarded and tested at X-band. The two-stage phase modulator schematic and assembly are illustrated in Figs. 8 and 9, respectively. The circuit layout is etched on a 0.508-mm-thick RT/Duroid 6002 soft substrate of dielectric constant $K = 2.94$. The constant 1.5 gamma varactor (Microwave Associates MA 46471-96) is mounted with a contact strip at the end of an optimized 30-ohm, 2.79-mm-wide transformer. The diode capacitance at -4 V is equal to 0.65 pF. TRACK 6.35-mm flange micropuck circulators (79*9001) and isolators (89*9001) are used in this breadboard. Their measured port-to-port one-way insertion loss is 1 ± 0.1 dB. The measured values of phase deviation and insertion loss for the phase modulator are compared with the predicted results in Fig. 10. The measurements were conducted by

using a test RF signal level of +9 dBm. The insertion loss for this unit is 8.3 ± 0.5 dB, the phase deviation is ± 150 deg, with a linearity better than ± 8 percent for a bias voltage range of 4.5 ± 2.5 V. The model accurately predicted a linear phase slope of 60 deg/V. The predicted insertion loss, including circulator and isolator losses, is 7.4 ± 0.4 dB, which is about 0.9 dB lower than the measured values. The discrepancy in the insertion loss probably resulted from the unaccounted-for interconnecting microstrip and radiation losses. The measured phase-delay variation is 0.252 psec/deg C over the design temperature range of -30 deg C to $+85$ deg C. The size of the breadboard modulator is $61 \times 36 \times 14.3$ mm (Fig. 9). The driver circuit is not included in this breadboard layout.

C. Phase-Modulator Op-Amp Driver Circuit

The phase-modulator driver circuit schematic is shown in Fig. 11. The functions of the phase-modulator drive circuit are to sum and amplify the modulation input signals and to provide composite drive voltage to the varactor diodes. The modulation input signals include the spacecraft telemetry, ranging, and differential one-way ranging (DOR) signals. The modulation frequency ranges from 1 kHz to 20 MHz. The selected wideband op-amp for this application is Comlinear CLC 505. The size of the modulator driver circuit is $38 \times 25 \times 14.3$ mm.

Figure 12 shows the predicted and measured gain versus frequency characteristics of the phase-modulator driver circuit at -35 deg C, $+25$ deg C, and $+85$ deg C. The measured response is flat and drops only 0.4 dB, from 0.5 MHz to 20 MHz. The measured 3-dB bandwidths at -35 deg C, $+25$ deg C, and $+85$ deg C are 92 MHz, 84 MHz, and 75 MHz, respectively. The measured results are actually better than the circuit response predicted by the Simulation Program with Integrated Circuit Emphasis (SPICE) model. The model correctly predicted higher gain at lower temperatures. The dc power consumption with an 8-volt peak-to-peak output swing is 112 mW; when no signal is applied, the power consumption is 32 mW.

IV. Conclusions

Single-section and two-section analog X-band reflection phase modulators were developed. The phase modulator performance is accurately predicted by the theoretical modeling and simulations. By using a given package-type abrupt-junction varactor diode and a circulator, a phase modulator was realized. From 8257 MHz to 8634 MHz, the measured voltage-controlled linear phase deviation, linearity tolerance, and insertion loss were ± 46 deg, ± 8 percent,

and 2 ± 0.3 dB, respectively, over the test temperature range of -20 deg C to $+75$ deg C. The static phase delay was found to be 0.16 psec/deg C. By using hyperabrupt junction varactor diodes of a constant gamma of 1.5, a full ± 2.5 radian deviation phase modulator was realized by cascading two phase-shifter circuits. From 8400 to 8500 MHz, the measured phase shift and insertion loss

are ± 2.5 radians with better than ± 8 percent linearity and 8.3 ± 0.5 dB, respectively. An efficient phase modulator driver circuit was also developed and tested. The measured gain-frequency characteristics over the design temperature range agreed well with the predicted results. The modulator driver 3-dB bandwidth and dc power consumption at 25 deg C are 84 MHz and 112 mW, respectively.

Acknowledgments

The authors gratefully acknowledge the valuable support and comments of A. W. Kermode during the course of this work. The authors also thank Mark Bockmann of Motorola (Strategic Electronics Division, Chandler, Arizona) for the phase-modulator breadboard implementation.

References

- [1] N. R. Mysoor, "A Low-Loss Linear Analog Phase Modulator for 8415 MHz Transponder Applications," *TDA Progress Report 42-96*, vol. October-December 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 172-178, February 15, 1989.
- [2] R. Garver, "360° Varactor Linear Phase Modulator," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-17, pp. 137-147, March 1969.
- [3] R. Garver, "Broad-Band Diode Phase Shifters," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-20, pp. 314-323, May 1972.

Table 1. Summary of predicted performance for multiple section phase modulators

Varactor gamma	Phase shift, deg	Linearity, percent	Insertion loss, dB	Insertion loss variation, dB	Sensitivity, deg/V	Nominal bias voltage, Vdc	Modulation swing, V	Number of sections
0.47	± 140	$< \pm 8$	5.35	± 0.15	60	4.5	± 2.5	3
1.25	± 143	$< \pm 8$	2.9	± 0.08	60	4.5	± 2.5	2
1.5	± 142	$< \pm 8$	2.55	± 0.1	60	4.5	± 2.5	2
Specification	± 143.5	$< \pm 8$	10 dB max	± 0.45	57	—	—	

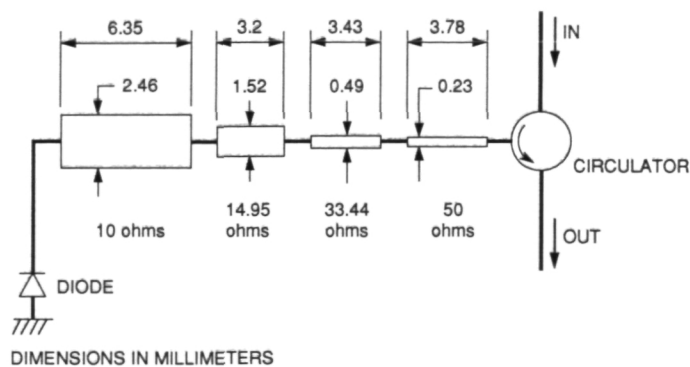


Fig. 1. Phase modulator circuit layout etched on a 0.254-mm-thick soft substrate dielectric constant $K = 10.5$.

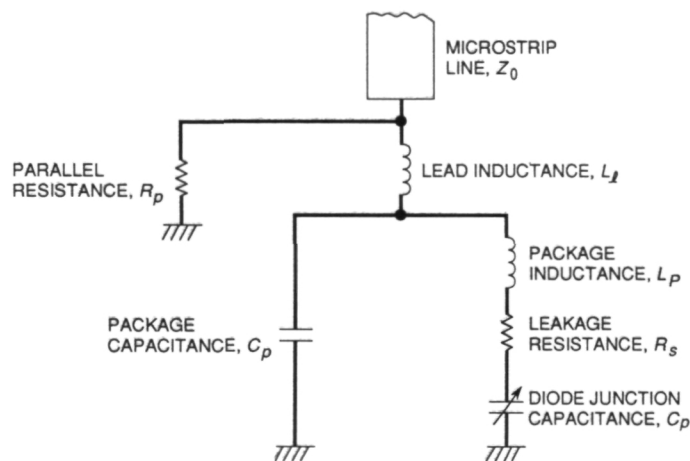


Fig. 2. Phase modulator circuit model.

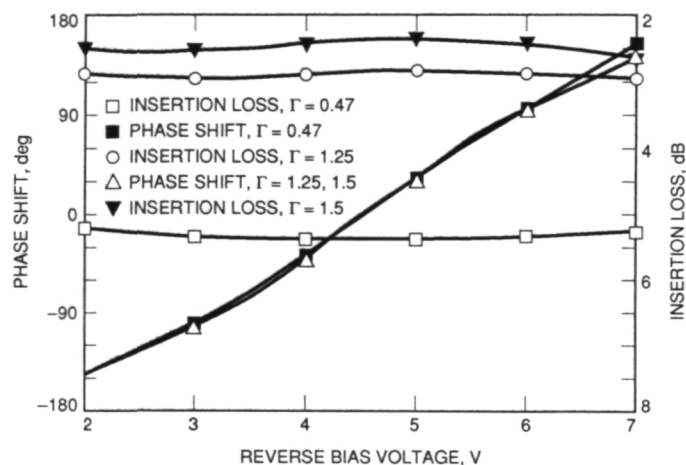


Fig. 4. Predicted phase shift and insertion loss for a two-section phase modulator with $\Gamma = 1.25, 1.5$, and for a three-section phase modulator with $\Gamma = 0.47$.

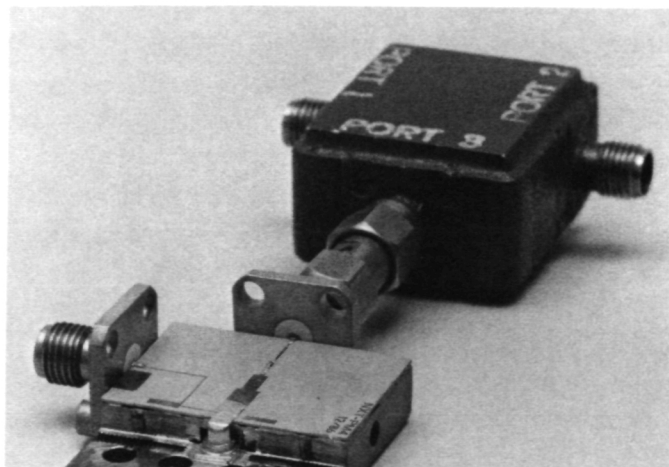


Fig. 5. X-band (8415 MHz) phase modulator.

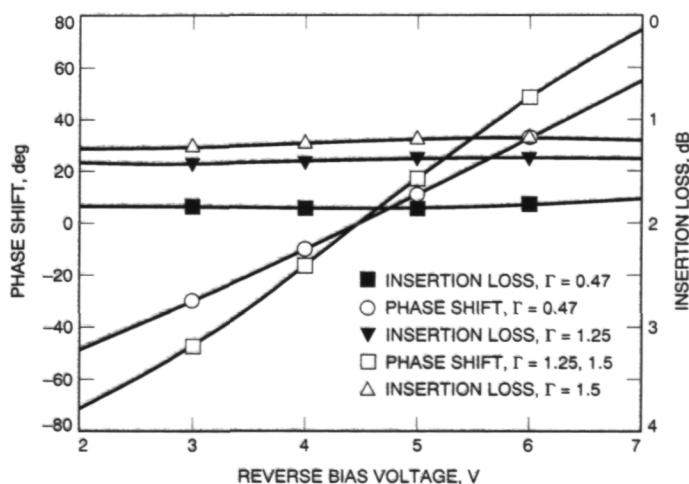


Fig. 3. Predicted phase shift and insertion loss as a function of reverse bias voltage for a single-section phase modulator for three different values of γ .

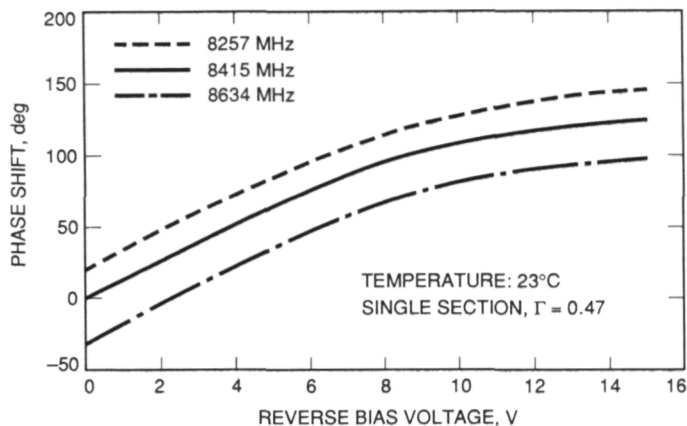


Fig. 6. Measured phase shift as a function of dc bias voltage and frequency.

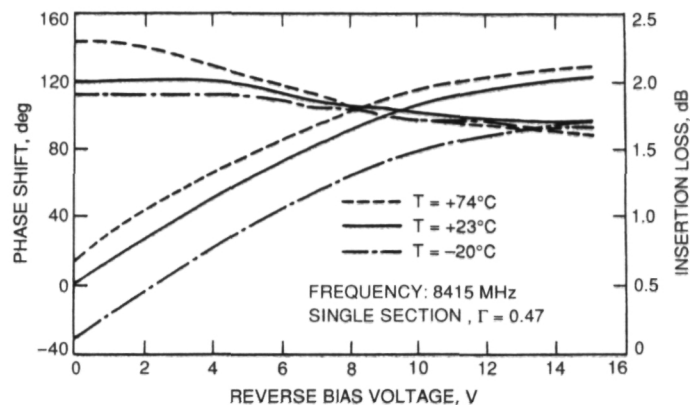


Fig. 7. Measured phase shift versus dc bias voltage and temperature.

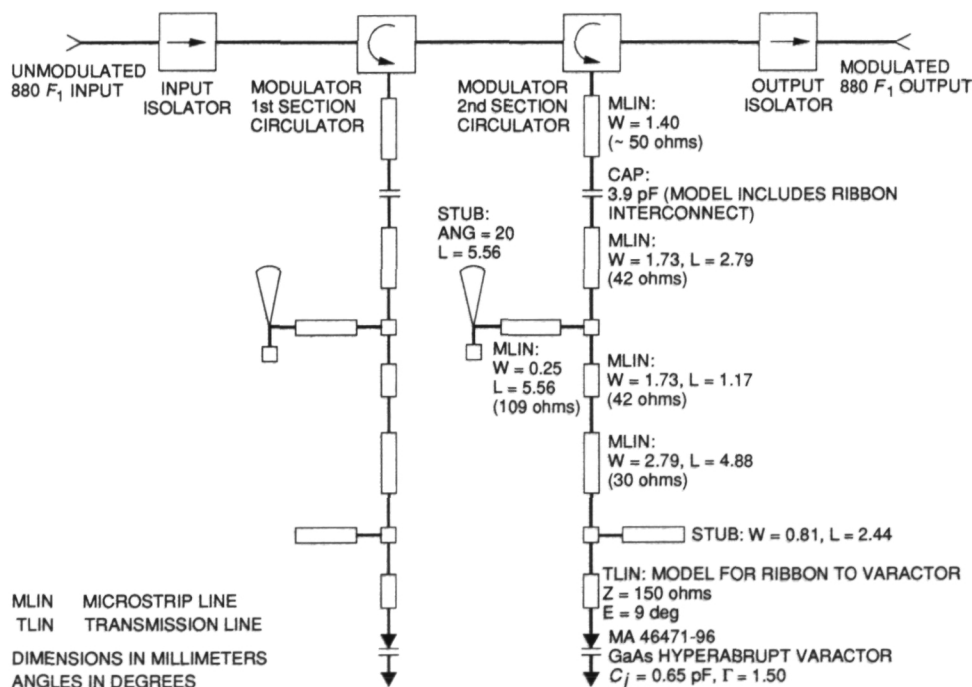


Fig. 8. Deep Space Spacecraft Transponder X-band (8415-MHz) phase modulator schematic.

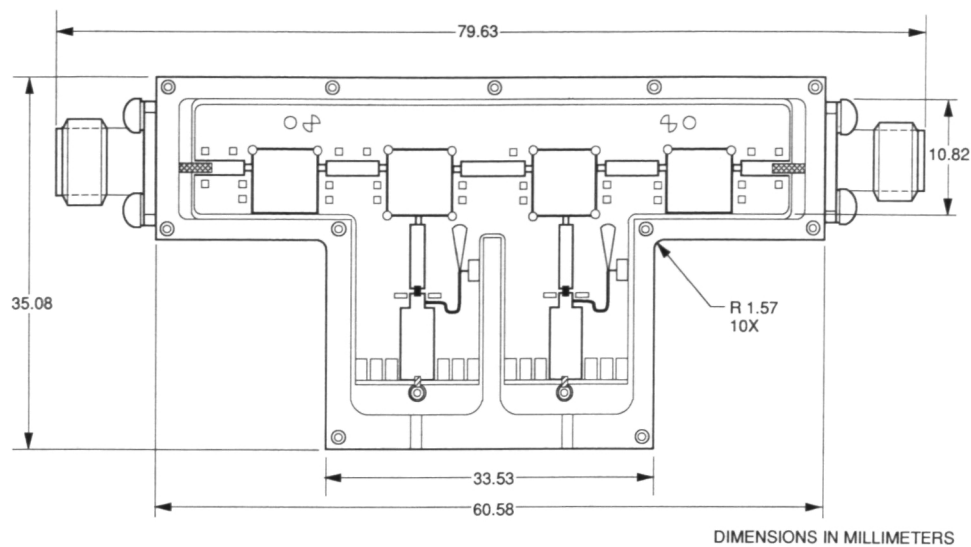


Fig. 9. Deep Space Spacecraft Transponder phase modulator assembly.

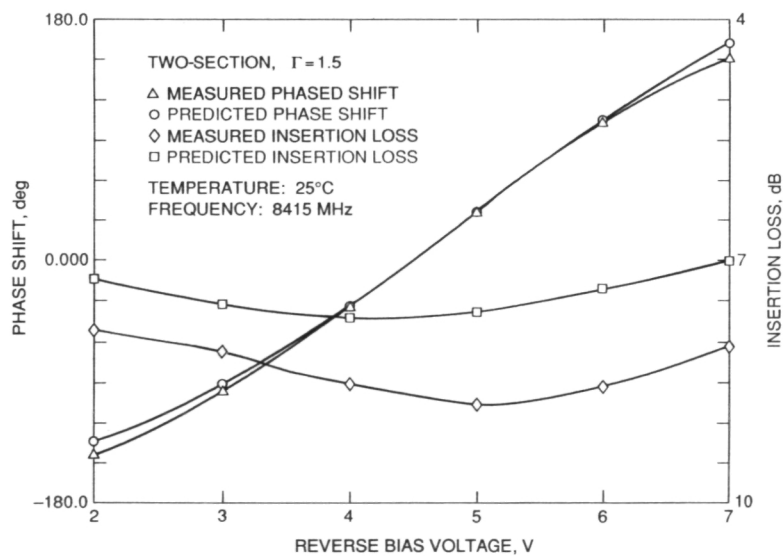


Fig. 10. Measured and predicted results for the two-section phase modulator.

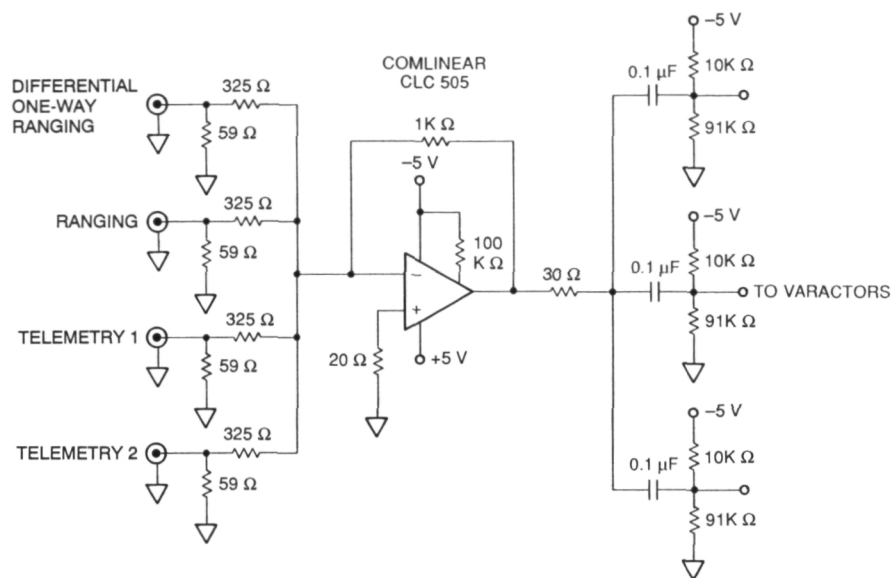


Fig. 11. Deep Space Spacecraft Transponder phase modulator driver circuit schematic.

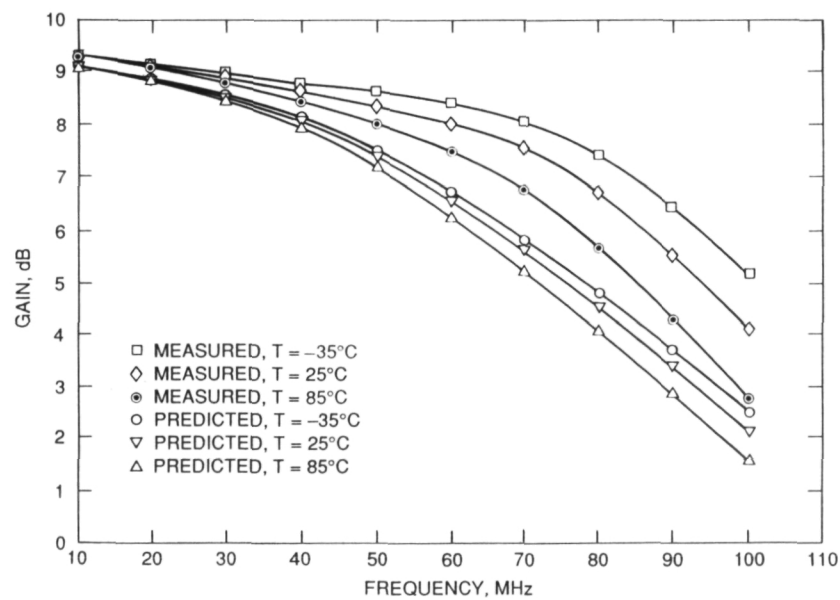


Fig. 12. Measured and predicted gain versus frequency for the phase modulator driver circuit.

Orbital Debris Radar Instrumentation

R. M. Goldstein

Telecommunications Science and Engineering Division

L. W. Randolph

TDA Engineering

In order to increase the usefulness of Goldstone orbital debris observations, several improvements are planned to the experimental Goldstone radar. The first improvement is to add a ranging capability of 1-km accuracy to the radar. This would provide much more detailed information about whatever debris is observed. The second improvement is to widen the bandwidth of the system from 10 kHz to 85 kHz so that particles in more elliptical orbits can be monitored. The two antennas could then also be pointed away from the zenith, say south, where orbits of lesser inclination might also be observed. In any case, the addition of ranging requires a substantial increase of system bandwidth. This article describes the instrumentation necessary to achieve this, together with required signal-processing modifications.

I. Introduction

Orbital debris has been recognized as a rapidly increasing hazard for United States and world activities in space. Individual pieces of debris larger than about 10 centimeters are now routinely tracked and catalogued by the Space Surveillance Network, operated by the Department of Defense [1]. Research in debris population modelling and shielding is being done at the National Aeronautics and Space Administration's (NASA's) Johnson Space Center [2]. In addition, Space Station planners have included enough shielding in their design to withstand impacts with one-centimeter and smaller sized objects. However, orbiting particles even below one millimeter in size can be lethal to an astronaut working outside of this protective shielding.

In an attempt to assess the magnitude of this hazard, two preliminary radar experiments have been performed: one at the Arecibo Observatory [3] and the other at Goldstone [4]. The results gave an irregular detection rate of about two events per hour. This corresponds to a particle flux of 6.4 objects per square km per day of equivalent size of 1.8 mm or greater.

Only a small amount of orbital parameter space was monitored by these experiments. For the Goldstone experiments, the particles were illuminated at X-band (8.5 GHz) by the transmitter and antenna at Deep Space Station (DSS) 14; the receiver was at DSS 13, 21.6 km away. The two antennas were aimed at a point midway between them, and 600 km above ground level. The intersection

of the two beams spanned an altitude difference of only 40 km. Due to bandwidth limitations of the radar system, only particles with near circular orbits could be observed. Any debris moving with a radial velocity greater than 88 m/sec would be Doppler-shifted outside of the receiver bandwidth.

II. Instrumentation Improvements

In order to increase the usefulness of Goldstone orbital debris observations, several improvements are planned to the experimental Goldstone radar. The first improvement is to add a ranging capability of 1-km accuracy to the radar. This would provide much more detailed information about whatever debris is observed. The second improvement is to widen the bandwidth of the system from 10 kHz to 85 kHz so that particles in more elliptical orbits can be monitored. The two antennas could then also be pointed away from the zenith, say south, where orbits of lesser inclination might also be observed. In any case, the addition of ranging requires a substantial increase of system bandwidth. The third improvement involves more efficient signal processing. The system is able to perform the bulk of the processing on line, simultaneously with the observations, so that at the end of a track the data are already reduced and recorded. Thus, transporting the many rolls of magnetic tape, and the long, overnight computer runs of the earlier Goldstone experiment will no longer be necessary.

III. System Design

Ranging systems in the Deep Space Network (DSN) have always required accurate knowledge of the range rate. In the case of orbital debris, however, such information is not available. Particles can be expected to intersect the beams with considerable variation of radial velocities. Weak signal detection requires a separate correlation for each range-Doppler cell of interest, a prohibitive work load for the usual pseudonoise modulation system.

The design uses, instead, a chirp signal, which can search out a linear combination of range-Doppler cells simultaneously. Since, at most, one object is expected to be in the beam at a time, alternate use of up- and down-chirps enables the estimation of both range and Doppler.

Transmission will be a linear chirp (up or down) of 50-kHz bandwidth and 2-msec duration; 6 msec of receive time will follow. The typical particle of debris will remain within the DSS 14 beam for about six of these eight millisecond cycles.

Reception consists of the digitization of the received signals at a rate of 85 kHz (both in-phase and quadrature), the convolution of the samples with the appropriate reference, and the summation of squared magnitudes for each set of six send-receive cycles. These calculations are to be done in real time, in parallel with the signal sampling. When any result of the above calculations exceeds a preset threshold, those data are saved for additional analysis. In this way, only the "interesting" results from an entire radar pass need to be considered further, and are saved on the computer floppy disk.

IV. Hardware

The system described above has been implemented on a standard personal computer (PC) clone with an AT bus and an Intel 386 processor. Plugged into one of the expansion slots of the PC is a signal-processing board purchased from the Spectrum Signal Processing Company. This board contains an AT&T DSP32C processor, capable of 25 million floating point operations per second; 32 kbytes of high-speed memory and 128 kbytes of somewhat slower memory; an analog interface consisting of a timer, two Burr-Brown analog-to-digital and two digital-to-analog converters capable of a maximum speed of 100,000 samples per second each; an interface to the host PC that supports both direct memory access to the DSP32C and interrupts to the PC; and a high-speed serial and parallel input/output port.

This board handles digitization of the received signals, generation of the transmitted waveform, and signal processing required for orbital debris detection. Most of the computations are done on the DSP32C, which can perform a 512-point complex fast Fourier transform in 1.6 msec. The host PC performs the file keeping and display functions.

V. Software

The host computer is programmed in Microsoft C. Several C-callable routines are used to upload and download data and programs between the host computer and the DSP32C processor. In addition, much use has been made of a debugging routine for the assembly language of the DSP32C. These routines were written by the Loughborough Company and were supplied with the signal-processing board.

Programs for the DSP32C were written on the host computer in assembly language, with the aid of an AT&T assembler-linker package. A C compiler was provided with

the DSP32C, but the resultant code was both too long and too slow.

Several programs have been written for the system, each with one part in C for the host computer and a second part in assembly language for the DSP32C processor. The first program is the operational program, which implements the radar system functions, has a display for the operators, and reduces the data to a floppy disk at the conclusion of a track. In addition, there are test programs to ensure that actual signals of the correct level are applied, and that the transmitter and receiver can operate in harmony in a closed loop.

VI. Configuration for Experiments

The above-described hardware and software are now installed at DSS 14. When antenna time for the radar becomes available, the new system will be tested using the DSS-14-DSS-13 combination as in the previous experiments. If at all possible, the receiver will be located at DSS 15. With DSS 14 transmitting and DSS 15 receiving, the beam intersection would form a column extending from about 200 km outward. Thus, altitudes from 200 km to 1000 km could be monitored instead of only the 40-km thick area that is presently available with the DSS-14-DSS-13 combination, resulting in much more efficient use of DSN antenna time for monitoring orbital debris.

References

- [1] Interagency Group (Space), *Report on Orbital Debris*, National Security Council, Washington, D.C., February 1989.
- [2] R. C. Reynolds and A. E. Potter, *Orbital Debris Research at NASA Johnson Space Center 1986-1988*, NASA Technical Memorandum 102155, Johnson Space Center, Houston, Texas, September 1989.
- [3] T. W. Thompson, R. M. Goldstein, D. B. Campbell, A. E. Stansbery, and A. E. Potter, "Radar Detection of Centimeter-Sized Orbital Debris, Preliminary Arecibo Observations at 12.5 cm Wavelength," submitted to *Geophysical Research Letters*.
- [4] R. M. Goldstein and L. W. Randolph, "Rings of Earth Detected by Orbital Debris Radar," *TDA Progress Report 42-101*, vol. January-March 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 191-195, May 15, 1990.

Errata

T. Y. Otoshi, S. R. Stewart, and M. M. Franco have submitted the following errata to "A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part I: Design and Ground Tests" that appeared in *The Telecommunications and Data Acquisition Progress Report 42-103*, vol. July–September 1990, pp. 135–150. The last sentence of the third paragraph on page 137 should be replaced with:

The corrected system temperature is obtained by multiplying each point by the corresponding linearity factor shown in Fig. 10(a) [5]. However, it can be seen that the corrected value would be different by about only 1 percent. It should be pointed out that in this article the above term "uncorrected" measured operating system noise temperature was used to refer to the system temperature that had already corrected for gain change, but not corrected for nonlinearity.

Errata

N. R. Mysoor, J. D. Perret, and A. W. Kermode have submitted the following errata to "Design Concepts and Performance of NASA X-Band (7162 MHz/8415 MHz) Transponder for Deep-Space Spacecraft Applications" that appeared in *The Telecommunications and Data Acquisition Progress Report 42-104*, vol. October-December 1990, pp. 247-256. Table 2 and Figure 1 should be replaced with the table and figure on the following pages.

Table 2. Transponder carrier phase tracking loop parameters

Transponder	Signal condition												
	Design			Threshold, $\alpha_0 = 0.0531$			10 dB above threshold, $\alpha_{10} = 0.1665$			Strong signal, 50 dB above threshold, $\alpha_S = 1.0$			
	K_v , sec ⁻¹	τ_1 , sec	τ_2 , sec	ζ_0	ω_{n0} , rad/sec	$2B_{L0}$, Hz	ζ_{10}	ω_{n10} , rad/sec	$2B_{L10}$, Hz	ζ_S	ω_{nS} , rad/sec	$2B_{LS}$, Hz	$\Delta\omega$, Hz/sec
NST	1.44×10^7	2910	0.0833	0.68	16.2	17.0	1.20	40.3	28.7	2.90	70.3	209.9	394
NST + XSDC	4.88×10^7	2910	0.0833	1.24	29.8	43.1	2.20	52.8	122.3	5.39	129.5	704.5	1334
GLL	1.62×10^7	3732	0.0423	0.32	15.2	16.7	0.57	27.1	26.9	1.42	66.0	105.3	347
GLL + XSDC	5.73×10^7	3732	0.0423	0.73	34.1	36.6	1.30	90.0	60.4	2.66	123.9	341.2	1222
MAG	1.44×10^7	728	0.0208	0.42	40.4	41.0	1.49	118.3	71.5	1.50	126.0	210.0	1263
MAG + XSDC	4.88×10^7	728	0.0208	0.62	59.7	61.6	1.10	105.6	140.1	2.69	258.9	721.2	5335
DST	2.2×10^7	3556	0.0556	0.50	18.0	18.0	0.88	31.9	37.2	2.17	78.1	178.5	550

K_v = DC gain of the PLL, sec⁻¹

τ_1 = Time constant associated with the open-loop pole (phase lag) of the loop filter, sec

τ_2 = Time constant associated with the open-loop zero (phase lead) of the loop filter, sec

$\alpha_0, \alpha_{10}, \alpha_S$ = Limiter suppression factor at threshold, 10 dB above threshold, and strong signal

$\zeta_0, \zeta_{10}, \zeta_S$ = PLL damping factor at threshold, 10 dB above threshold, and strong signal

$\omega_{n0}, \omega_{n10}, \omega_{nS}$ = PLL natural frequency at threshold, 10 dB above threshold, and strong signal, rad/sec

$2B_{L0}, 2B_{L10}, 2B_{LS}$ = PLL noise-equivalent bandwidth at threshold, 10 dB above threshold, and strong signal, Hz

$\Delta\omega$ = PLL acquisition and tracking rate at strong signal, Hz/sec

NST = NASA Standard DST (S-band)

GLL = Galileo transponder (S-band)

MAG = Magellan transponder (S-band)

DST = Deep space transponder (X-band design)

XSDC = External X-band to S-band downconverter

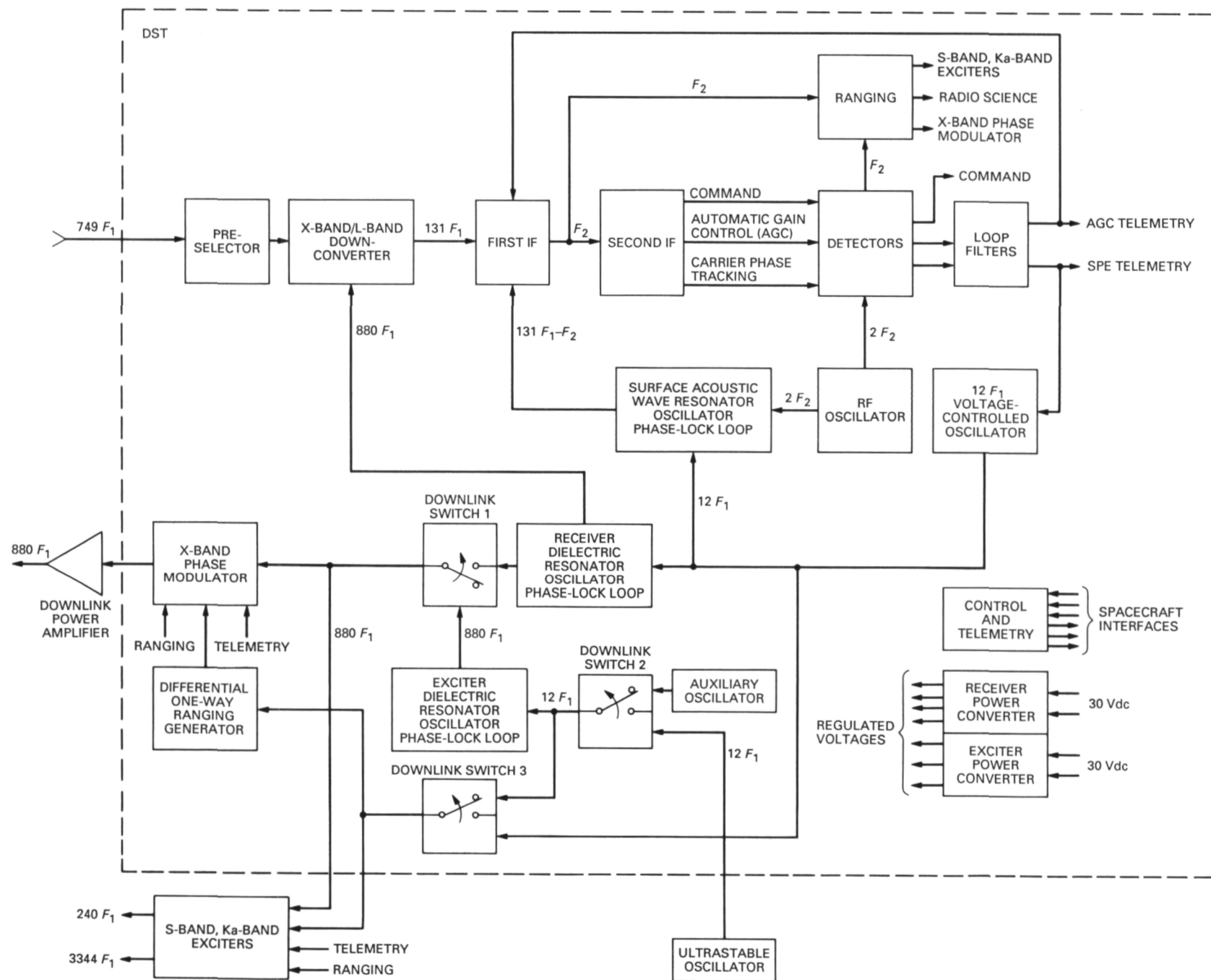


Fig. 1. Deep-space transponder.

1. Report No. 42-105	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle The Telecommunications and Data Acquisition Progress Report		5. Report Date May 15, 1991	
		6. Performing Organization Code	
7. Author(s) Editor: E. C. Posner		8. Performing Organization Report No. 42-105	
9. Performing Organization Name and Address JET PROPULSION LABORATORY California Institute of Technology 4800 Oak Grove Drive Pasadena, California 91109		10. Work Unit No.	
		11. Contract or Grant No. NAS7-918	
		13. Type of Report and Period Covered TDA Progress Report, vol. January-March 1991	
12. Sponsoring Agency Name and Address NATIONAL AERONAUTICS AND SPACE ADMINISTRATION Washington, D.C. 20546		14. Sponsoring Agency Code	
15. Supplementary Notes			
16. Abstract <p>This quarterly publication provides archival reports on developments in programs managed by JPL's Office of Telecommunications and Data Acquisition (TDA). In space communications, radio navigation, radio science, and ground-based radio and radar astronomy, it reports on activities of the Deep Space Network (DSN) in planning, supporting research and technology, implementation, and operations. Also included are standards activity at JPL for space data and information systems and reimbursable DSN work performed for other space agencies through NASA. The preceding work is all performed for NASA's Office of Space Communications (OSC). The TDA Office also performs work funded by two other NASA program offices through and with the cooperation of OSC. These are the Orbital Debris Radar Program with the Office of Space Systems Development and 21st Century Communication Studies (with the Office of Exploration).</p> <p>For the High-Resolution Microwave Survey (HRMS), <i>The TDA Progress Report</i> reports on implementation and operations for searching the microwave spectrum. In solar system radar, it reports on the uses of the Goldstone Solar System Radar for scientific exploration of the planets, their rings and satellites, asteroids, and comets. In radio astronomy, the areas of support include spectroscopy, very long baseline interferometry, and astrometry. These three programs are performed for NASA's Office of Space Science and Applications (OSSA), with OSC funding DSN operational support.</p> <p>Finally, tasks funded under the JPL Director's Discretionary Fund and the Caltech President's Fund that involve the TDA Office are included.</p>			
17. Key Words (Selected by Author(s)) Astronautics Communications Mathematical and Computer Sciences Space Sciences		18. Distribution Statement Unlimited/Unclassified	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 152	22. Price